

A-59941-4

1 / 45



FIG. 1B



FIG. 1A

+

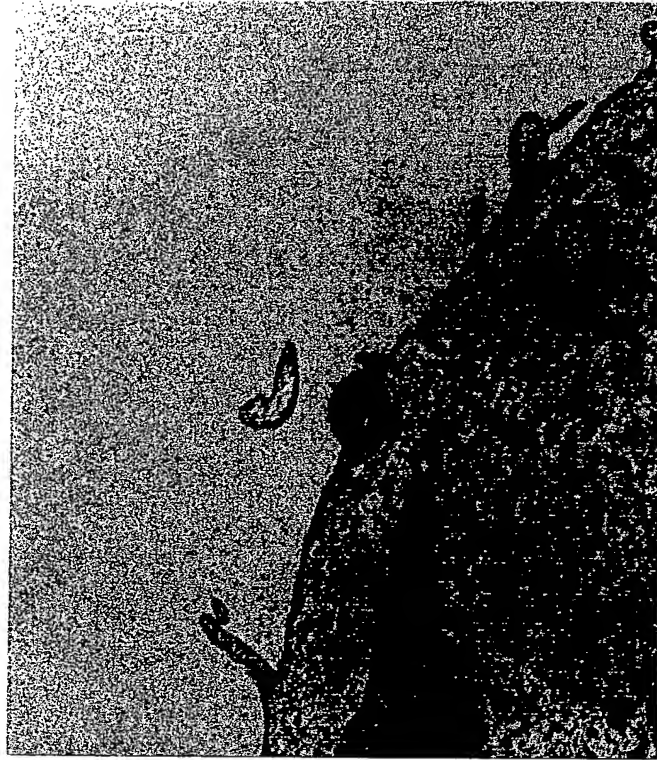


FIG._2B



FIG._2A

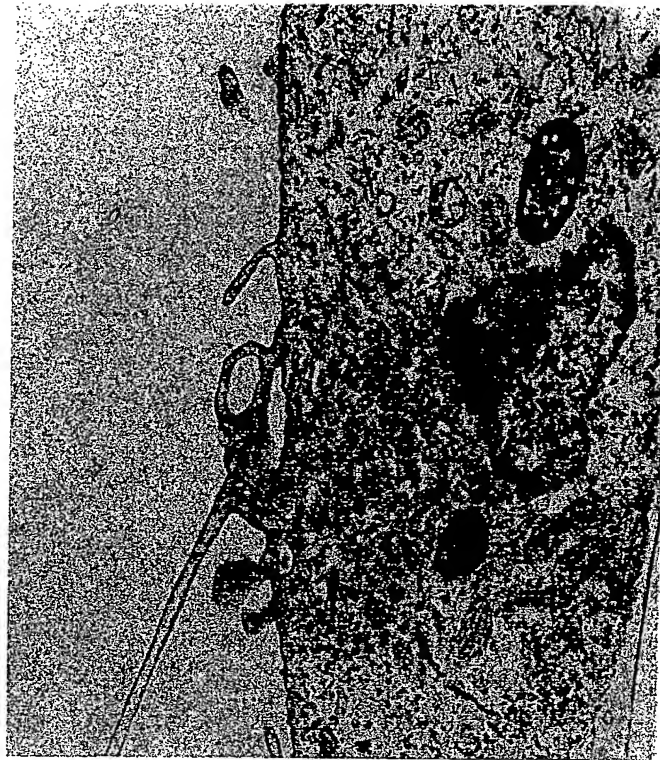


FIG._2D

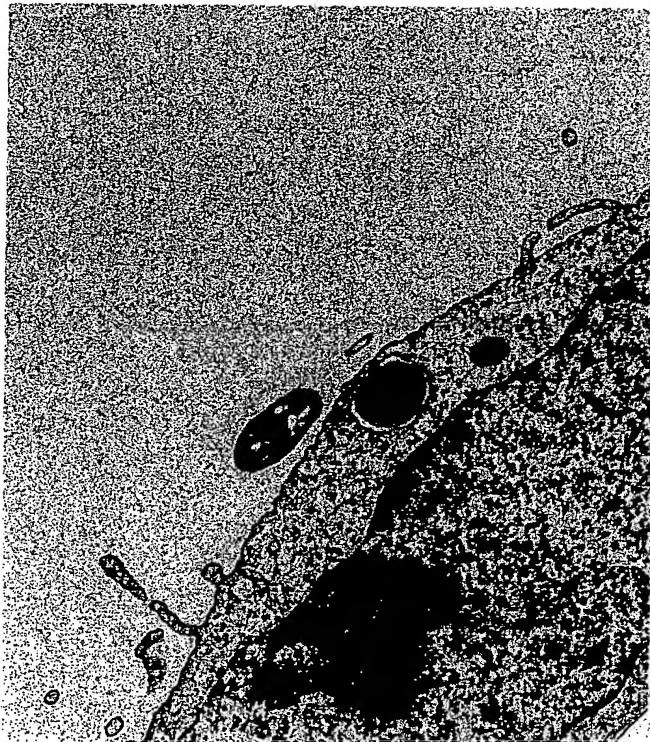


FIG._2C

FIG._3

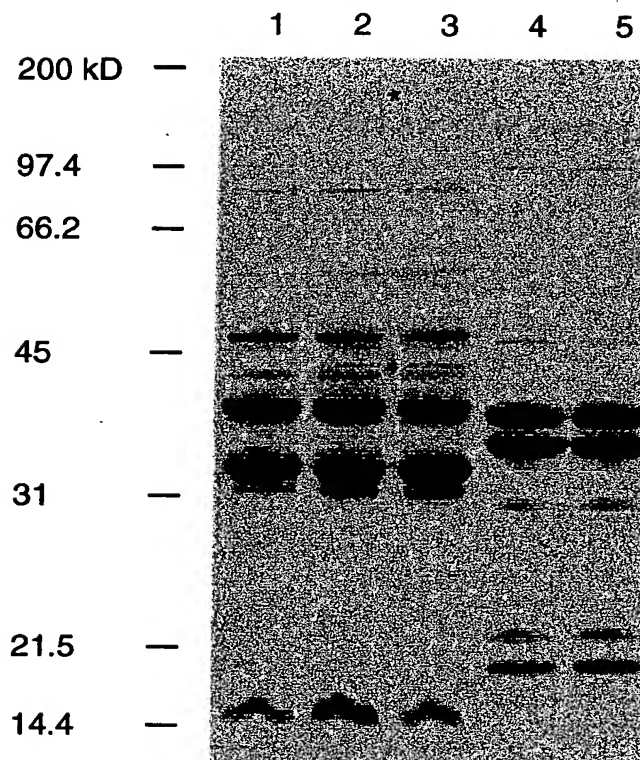
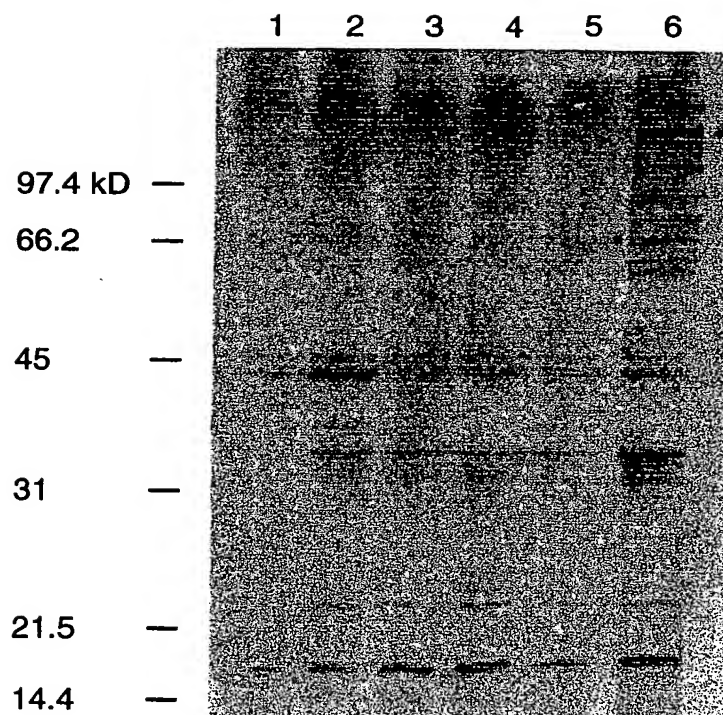
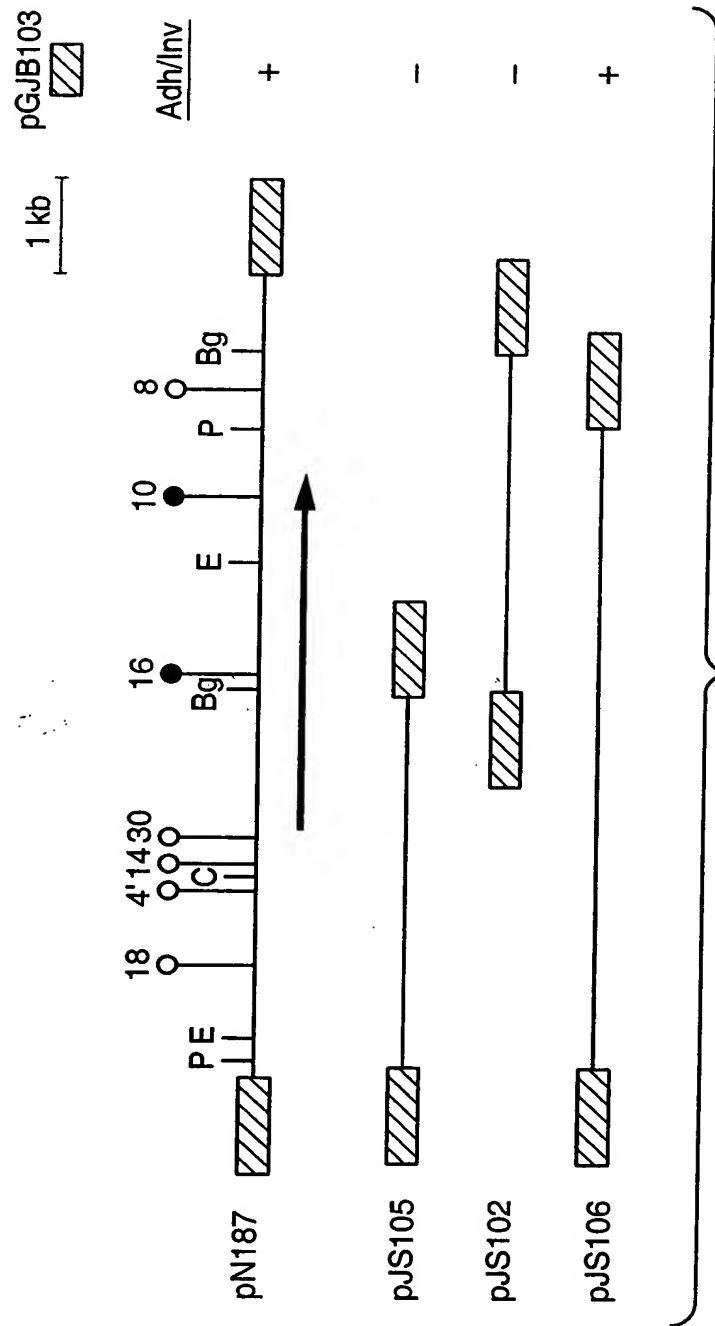


FIG._5



**FIG. 4**

```

10      30      50      70      90
TCAATAGTCGTTAACTAGTATTTTAAATACGAAAAATTAATTAATAAACATTATGAAAAAAGCTGATTTTCGTCCTTAATTTT
-35      -10
110      130      150      170
TTAAACCGCTTGCAATTCATTAGGGATAGTATCGCAAGCGTGGGCTGGTCACACTTATTTTGGGATTGATTACCAATATTATTCGTGATTTT
L T A C I S L G I V S Q A W A G H T Y F G I D Y Q Y Y R D F
190      210      230      250      270
GCCGAGATAAAGGGAAGTTCACAGTTGGGGCTCAAAATATTAAGGTTTATTAACAACAAGGCAATTAGTTGGCACATCAATGACAAAA
A E N K G K F T V G A Q N I K V Y N K Q G Q L V G T S M T K
290      310      330      350
GCCCCGATGATTGATTTTCTGTAGTGCACGTAACGGCGTGGCAGCCTTGGTTGAAAAATCAATATATTGTGAGCGTGGCACATAACGTA
A P M I D F S V V S R N G V A A L V E N Q Y I V S V A H N V
370      390      410      430      450
GGATATACAGATGTTGATTTTGGTGCAGAGGGAAACAACCCGATCAACATCGTTTACTTATAAGATTGTAAACGAAATAACTACAAA
G Y T D V D F G A E G N N P D Q H R F T Y K I V K R N N Y K
470      490      510      530
AAAGATAATTTACATCCTTATGAGGACGATTACCATATAATCCACGATTACATAAAATTCGTTACAGAAAGCGGCTCCAATTGATATGACTTCG
K D N L H P Y E D D Y H N P R L H K F V T E A A P I D M T S
550      570      590      610      630
AATATGAATGGCAGTACTTATTCAGATAGAACAAAAATATCCAGAACGTGTCGTATCGGCTCTGGACGGCAGTTTGGCGAAATGATCAA
N M N G S T Y S D R T K Y P E R V R I G S G R Q F W R N D Q
650      670      690      710
GACAAAGCGCACAAGTTGCCGGTGCATATCATTTATCTGACAGCTGGCAATACACACAATCAGCGTGGAGCAGGTAATGGATATTCGTAT
D K G D Q V A G A Y H Y L T A G N T H N Q R G A G N G Y S Y

```

FIG.-6A

730 750 770 790 810
 TTGGGAGCGCATGTTCTGTAAGCGGGAGAAATATGGTCCATTACCGATTGCAGGCTCAAAGGGGACAGTGGTCTCCGATGTTTATTTAT
 L G G D V R K A G E Y G P L P I A G S K G D S G S P M F I Y

830 850 870 890
 GATGCTGAAAAACAATAATGGTTAATTAATGGGATATTACGGGAAGGCAACCCCTTTTGAAGGCAAGAAATGGGTTTCAATTTGGTTTCGC
 D A E K Q K W L I N G I L R E G N P F E G K E N G F Q L V R

910 930 950 970 990
 AAATCTTATTTTGATGAAATTTTCGAAAGAGATTACATACATCATTACACCCGAGCTGGTAATGGAGTGACACAAATTAGTGGAAT
 K S Y F D E I F E R D L H T S L Y T R A G N G V Y T I S G N

1010 1030 1050 1070
 GATAATGGTCAGGGGCTATAAATCAGAAATCAGGAATACCATCAGAAATTAATAATACGTTAGCAAAATATGAGTTTACCTTTGAAAAGAG
 D N G Q G S I T Q K S G I P S E I K I T L A N M S L P L K E

1090 1110 1130 1150 1170
 AAGGATAAAGTTCAATAATCCTAGATATGACGGACCTAATATTATTCTCCACGTTTAAACAATGGAGAAACGCTATATTTTATGGATCAA
 K D K V H N P R Y D G P N I Y S P R L N N G E T L Y F M D Q

1190 1210 1230 1250
 AAACAAGGATCATTAATCTTCGCAATCTGACATTAACCAAGGGCGGTGCTCTTTATTGAGGGTAATTTTACAGTATCTCCAAATCT
 K Q G S L I F A S D I N Q G A G G L Y F E G N F T V S P N S

1270 1290 1310 1330 1350
 AACCAAACTTGGCAAGGAGCTGGCATACATGTAAGTGAATAAGTACACCGTTACTTGGAAAGTAAATGGCGTGGAAACATGATCGACTTCT
 N Q T W Q G A G I H V S E N S T V T W K V N G V E H D R L S

1370 1390 1410 1430
 AAAATTGGTAAAGAACATTGCACGTTCAAGCCAAAGGGGAAATAAAGTTTCGATCAGCGTAGCGGATGGTAAAGTCAATTTTGGAGCAG
 K I G K G T L H V Q A K G E N K G S I S V G D G K V I L E Q

FIG.-6B

1450 1470 1490 1510 1530
 CAGGCACGATCAAGGCAACAACAGCCCTTTAGTGAAATTGGCTTGGTAGCGGCAGAGGACTGTTCAATTAAACGATGATAAACAA
 Q A D D Q G N K Q A F S E I G L V S G R G T V Q L N D D K Q

1550 1570 1590 1610
 TTTGATACCGATAAATTTTATTTTCGGCTTTCGGTGGTTCGCTTAGATCTTAAACGGCATTCATTAAACCTTTAAACGTATCCAAAATACG
 F D T D K F Y F G F R G G R L D L N G H S L T F K R I Q N T

1630 1650 1670 1690 1710
 GACGAGGGGCAATGATTGTGAACCATATAACAACCTCAAGCCGCTAATGTCACTATTACTGGGAACGAAAGCATTTGTTACCTAATGGA
 D E G A M I V N H N T T Q A A N V T I T G N E S I V L P N G

1730 1750 1770 1790
 AATAATATAATAACTTGATTACAGAAAAGAAATTGCCCTACAACGGTTGGTTGGCGAAACAGATAAAAATAACACAATGGCGGATTA
 N N I N K L D Y R K E I A Y N G W F G E T D K N K H N G R L

1810 1830 1850 1870 1890
 AACCTTATTTATAAACCAACCAAGATCGTACTTTGCTACTTTTCAGGTGGTGACAAATTTAAAGGCGGATATTACCCAAACAAAGGT
 N L I Y K P T T E D R T L L L S G G T N L K G D I T Q T K G

1910 1930 1950 1970
 AAACATATTTTCAGCGGTAGACCGACCGCACCGCCCTACAATCATTTAAATAAACGTTGGTCAGAAAATGGAAGGTATACCACAAGGCGAA
 K L F F S G R P T P H A Y N H L N K R W S E M E G I P Q G E

1990 2010 2030 2050 2070
 ATTGTGTGGGATCACGATTGGATCAACCGTACATTTAAAGCTGAAAACCTCCAAAATTAAGCGGAAAGTGCGGTGGTTCTCGCAATGTT
 I V W D H D W I N R T F K A E N F Q I K G G S A V V S R N V

2090 2110 2130 2150
 TCTTCAATTGAGGGAATTGGACAGTCAGCAATAATGCAAAATGCCACATTTGGTGTGTGTGCCAAAATCAACAAAATACCATTGACACGCGT
 S S I E G N W T V S N N A N A T F G V V P N Q Q N T I C T R

FIG.-6C

2170 TCAGATTGGACAGGATTAAACGACTTGTCAAAAAGTGGATTAAACCGATACAAAAGTTAATTCTATACCAAAAACAAATCAATGGC 2210 2230 2250
 S D W T G L T T C Q K V D L T D T K V I N S I P K T Q I N G
 2270 TCTATTAAATTTAACTGATAATGCAACGGCGAATGTTAAAGGTTTAGCAAAACTTAATGGCAATGTCACTTTAACAAATCACAGCCAATTT 2310 2330
 S I N L T D N A T A N V K G L A K L N G N V T L T N H S Q F
 2350 ACATTAAGCAACAATGCCACCCAAATAGGCAATATTCGACTTTCCGACAATTCAACTGCAACGGTGGATAATGCAAACTTGAACGGTAAT 2410 2430
 T L S N N A T Q I G N I R L S D N S T A T V D N A N L N G N
 2450 GTGCATTTAACGGATTCAAGCTCAATTTCTTTAAAAAACAGCCATTTTCGCACCAAAATTCAGGGAGACAAAAGGCACACAGTGACGTTG 2510
 V H L T D S A Q F S L K N S H F S H Q I Q G D K G T T V T L
 2530 GAAAATGCGACTTGGACAAATGCCCTAGCGATACATTCAGAAATTTAAACGCTAAATAACAGTACGATCACGTTAAATTCAGCTTATTCA 2590 2610
 E N A T W T M P S D T T L Q N L T L N N S T I T L N S A Y S
 2630 GCTAGCTCAAACAATACGCCACGTCGCCGTTTCATTAGAGACGGAAACAACGCCAACATCGGCAGAACATCGTTTCAACACATTGACAGTA 2670 2690
 A S S N N T P R R R S L E T E T P T S A E H R F N T L T V
 2710 AATGGTAAATTGAGTGGGCAAGGCACATTCCTCAATTTACTTCATCTTTATTTGGCTATAAAAAGCGATAAATTAAATTTATCCAAATGACGCT 2770 2790
 N G K L S G Q G T F Q F T S S L F G Y K S D K L K L S N D A
 2810 GAGGGCGATTACATATCTGTTCGCAACACAGGCAAGAAACCCGAAACCCCTTGAGCAATTAACCTTGGTTGAAAGCAAGATAATCAA 2850 2870
 E G D Y I L S V R N T G K E P E T L E Q L T L V E S K D N Q

FIG.-6D

2890 2910 2930 2950 2970
 CCGTTATCAGATAAGCTCAAAATTACTTTAGAAAATGACCAACGTTGATGCAGGTGCATTACGTTATAAATTAGTGAAGATGATGGCGAA
 P L S D K L K F T L E N D H V D A G A L R Y K L V K N D G E

 2990 3010 3030 3050
 TTCCGGCTTGCAATAACCAATAAAGAGCAGGAATTGCACAATGATTAGTAAGAGCAGAGCAAGCAGAACGAACATTAGAACCAACAA
 F R L H N P I K E Q E L H N D L V R A E Q A E R T L E A K Q

 3070 3090 3110 3130 3150
 GTTGAACCGACTGCTAAACACAAACAGGTGAGCCAAAAGTGCAGTCAAGAGAGCAGCGAGAGCAGCGTTTCCCTGATACCCCTGCCCTGAT
 V E P T A K T Q T G E P K V R S R R A A R A A F P D T L P D

 3170 3190 3210 3230
 CAAAGCCTGTAAACGCAATTAGAAAGCCAAACAAAGCTGACTGCTGAAACACAAAAGTAAGGCAAAAACAAAAGTGCGGTCA
 Q S L L N A L E A K Q A E L T A E T Q K S K A K T K K V R S

 3250 3270 3290 3310 3330
 AAAAGAGCAGTGTCTTCTGATCCCTGCTGATCAAAAGCCTGTTCGCATAGAAAGCCGCACCTTGAGGTTATTGATGCCCCACAGCAATCG
 K R A V F S D P L L D Q S L F A L E A A L E V I D A P Q Q S

 3350 3370 3390 3410
 GAAAAAGATCGTCTAGCTCAAGAGAAGCGGAAACAAACGCAAAACAAAGACTTGATCAGCCGTTATTCAAAATAGTGCGTTATCAGAA
 E K D R L A Q E E A E K Q R K Q K D L I S R Y S N S A L S E

 3430 3450 3470 3490 3510
 TTATCTGCAACAGTAAATAGTATGCTTTCTGTTCAGATGAATTAGATCGTCTTTTGTAGATCAAGCACAATCTGCCGTGTGGACAAAT
 L S A T V N S M L S V Q D E L D R L F V D Q A Q S A V W T N

 3530 3550 3570 3590
 ATCGCACAGGATAAAGACGCTATGATTCTGATGCGTTCCGTGCTTATCAGCAGCAGAAAACGAACTTACGTCAAATGGGGTGCAAAAA
 I A Q D K R R Y D S D A F R A Y Q Q Q K T N L R Q I G V Q K

10 / 45

FIG.-6E

3610 3630 3650 3670 3690
 GCCTTAGCTAATGGACGAATTGGGGCAGTTTCTCGCATAGCCGTTTCAGATAATACCTTTGATGAACAGGTTAAAAATCACGCGACATTA
 A L A N G R I G A V F S H S R S D N T F D E Q V K N H A T L

3710 3730 3750 3770
 ACGATGATGTCGGGTTTGGCCCAATATCAATGGGGCGATTACAAATTGGTGTAACGTGGGAACGGGAATCAGTGCAGTAAAAATGGCT
 T M M S G F A Q Y Q W G D L Q F G V N V G T G I S A S K M A

3790 3810 3830 3850 3870
 GAAGAACAAAGCCGAAAAATTCAATCGAAAAAGCGATAAATTATGGCGTGAATGCAAGTTATCAGTTCCTGTTAGGGCAATTGGGCATTTCAG
 E E Q S R K I H R K A I N Y G V N A S Y Q F R L G Q L G I Q

3890 3910 3930 3950
 CCTTATTTTGGAGTTAATCGCTATTTTATGTAACGTGAAAAATTATCAATCTGAGGAAGTGAGAGTGAAAAACGCCCTAGCCTTGCAATTAAAT
 P Y F G V N R Y F I E R E N Y Q S E E V R V K T P S L A F N

3970 3990 4010 4030 4050
 CGCTATAATGCTGGCATTCGAGTTGATATACATTTACTCCGACAGATAATATCAGCGTTAAGCCTTATTTCTTCGTCAATTATGTTGAT
 R Y N A G I R V D Y T F T P T D N I S V K P Y F F V N Y V D

4070 4090 4110 4130
 GTTTCAAACGCTAACGTACAAACACCGGTAAATCTCACGGTGTTCGCAACAACCATTTGGACGTTATTGGCAAAAAGAGTGGGATTAAAG
 V S N A N V Q T T V N L T V L Q Q Q P F G R Y W Q K E V G L K

4150 4170 4190 4210 4230
 GCAGAAATTTTACATTTCCAAATTTCCGCTTTTATCTCAAAATCTCAAGGTTCACTCAAACTCGGCAACAGCAAAATGTGGGCGTGAAATTG
 A E I L H F Q I S A F I S K S Q G S Q L G K Q Q N V G V K L

4250 4270 4290 4310
 GGCATCGTTGGTAAAAATCAACATAATTTTATCGTTTATTGATAAAACAAGGTGGGTGAGATCCACCTTTTATTTATTTCCAAATAAT
 G Y R W *

FIG.-6F

Hap	1	50
HK368IGA	MLNKKFRLNF	LTACISLGIV
HK393IGA	MLNKKFKLNF	IALTVAYALT
HK715IGA	MLNKKFKLNF	IALTVAYALT
HK61IGA	MLNKKFKLNF	IALTVAYALT
Consensus	M-----F-LNF	-----A-----
	51	100
Hap	AQNIKVYNKQ	GQLVGTSMTK
HK368IGA	ATNVLVKDKN	NKDLGTALPN
HK393IGA	ATNVEVRDKN	NRPLGNVLPN
HK715IGA	ATNVEVRDKN	NHSLGNVLPN
HK61IGA	ATNVEVRDKN	NQSLGSALPN
Consensus	A-N--V--K-	-----G-----
	101	150
HapNVGY	TDVDFGAEGN
HK368IGA	VSNVGSSELHF	GNLNGNMNNG
HK393IGA	VSNVGSSELHF	GNLNGNMNNG
HK715IGA	VSNVGSSELHF	GNLNGNMNNG
HK61IGA	VSNVGSSELHF	GNLNGNMNNG
Consensus	-----	-----
	151	200
Hap	...KKDNLH	PYEDDYHNPR
HK368IGA	TVTTEDQ.TQ	KRREDYIMPR
HK393IGA	AVTTEDQ.AQ	KRREDYIMPR
HK715IGA	AVTTEDQ.TQ	KRREDYIMPR
HK61IGA	FTTKEEQDAQ	KRREDYIMPR
Consensus	-----	-----DY--PR

FIG. 7A

201	ERVRIGSGRQ F.....WRNDQ	250	DKGDQVAGAY
Hap	AFVRLGSGSQ FIYKKGDNYS LIL.....N	NH.....EVGG		NNLKLVGDAY
HK368IGA	YFVRLGSGTQ FIYENGTRYE LWL.....G	KEGQKSDAGG		YNLKLVGDAY
HK393IGA	AFVRLGSGSQ FIYKKGDNYS LIL.....N	NH.....EVGG		NNLKLVGDAY
HK715IGA	AFVRLGSGSQ FIYKKGSRQY LILTEKDKQG	NLLRNWDVGG		DNLELVGNAY
HK61IGA	--VR-GSG-Q F-----	-----		-----V--AY
Consensus				
300				
251	HYLTAGNTHN ORGAGNGYSY LGG.....D	VRKAGEYGPL	300	PIAGSKGDSG
Hap	TYGIAAGTPYK VNHENNGLIG FGNSKEEHS D	PKGILSQDPL		TNYAVLGDSDG
HK368IGA	TYGIAAGTPYE VNHENDGLIG FGNSNNEYIN	PKEILSKKPL		TNYAVLGDSDG
HK393IGA	TYGIAAGTPYK VNHENNGLIG FGNSKEEHS D	PKGILSQDPL		TNYAVLGDSDG
HK715IGA	TYGIAAGTPYK VNHENNGLIG FGNSKEEHS D	PKGILSQDPL		TNYAVLGDSDG
HK61IGA	-Y--AG-----G-----	-----PL		-----GDSDG
Consensus				
350				
301	SPMFIYDAEK QKWLINGILR EGNPFEGKEN	GFQLVRKSYF	350	D.EIFERDLH
Hap	SPLFVYDREK GKWLFLGSYD FWAGYN....KKSQ		EWNIIKSYQFT
HK368IGA	SPLFVYDREK GKWLFLGSYD YWAGYN....KKSQ		EWNIIKPEFA
HK393IGA	SPLFVYDREK GKWLFLGSYD FWAGYN....KKSQ		EWNIIKPEFA
HK715IGA	SPLFVYDREK GKWLFLGSYD FWAGYN....KKSQ		EWNIIKHEFA
HK61IGA	SP-F-YD-EK -KWL--G--- --	-----KS--		-----I-----
Consensus				
400				
351	TSLYTRAGNG VYTISGNDNG QGSITQKSGI	PSEIKITLAN	400	MSLPLKEKDK
Hap	KDVLNKDSAG SLIGSKTDYS WSSNGKTSTI	TGGEK....S		LNVDLAD...
HK368IGA	EKIYEAYSAG SLIGSKTDYS WSSNGKTSTI	TGGEK....S		LNVDLAD...
HK393IGA	KTVLDKDTAG SLTGSNTQYN WNPTGKTSVI	SNGSE....S		LNVDLFD...
HK715IGA	EKIYQQYSAG SLTGSNTQYT WQATGSTSTI	TGGGE....P		LSVDLTD...
HK61IGA	-----G-----S-----	-----S-I		-----L-----
Consensus				

FIG..7B

Hap	401	VHNPRYDGP	N	I	Y	S	P	R	L	N	N	G	E	T	L	Y	F	M	D	Q	K	Q	G	S	L	I	F	A	S	D	I	N	Q	G	A	G	L	Y	F	E	G	N		450	
HK368IGA	GKD.																																											
HK393IGA	GKD.																																											
HK715IGA	SSQD																																											
HK61IGA	GKD.																																											
Consensus		-----																																											
Hap	451	FTVSPNSNQ.																																											
HK368IGA		YEVKGTSDNT																																											
HK393IGA		YEVKGTSDNT																																											
HK715IGA		YEVKGTSDST																																											
HK61IGA		YEVKGTSDST																																											
Consensus		--V---S---																																											
Hap	501	GENKGSISVG																																											
HK368IGA		GDNKGSLKVG																																											
HK393IGA		GDNKGSLKVG																																											
HK715IGA		GENKGSILKVG																																											
HK61IGA		GKNEGLLKVG																																											
Consensus		G-N-G---VG																																											
Hap	551	TDKFFYFGFRG																																											
HK368IGA		PNSIYFGFRG																																											
HK393IGA		PNSIYFGFRG																																											
HK715IGA		PNSIYFGFRG																																											
HK61IGA		PNSIYFGFRG																																											
Consensus		-----YFGFRG																																											
Hap	600	TDKFFYFGFRG																																											
HK368IGA		PNSIYFGFRG																																											
HK393IGA		PNSIYFGFRG																																											
HK715IGA		PNSIYFGFRG																																											
HK61IGA		PNSIYFGFRG																																											
Consensus		-----YFGFRG																																											

FIG._7C

601	ESIVLPNG...	NPYAFRRIKD	GGQLYLNLN	YTYVALRKGA	650
Hap	SLITDPNTIT	PYNIDAPDED	NPYAFRRIKD	GGQLYLNLN	YTYVALRKGA	
HK368IGA	NLITDPNNVS	IYVVKPLEDD	NPYAIRQIKY	GYQLYFNEEN	RTYVALKKDA	
HK393IGA	SLITDPNTIT	PYNIDAPDED	NPYAFRRIKD	GGQLYLNLN	YTYVALRKGA	
HK715IGA	SLITNPNTIT	SYNIEAQDDD	HPLRIRSIPI	R.QLYFNQDN	RSYYTLKKGA	
HK61IGA	--I--PN---	-----	-----	-----	-----	-----	
Consensus							
651	700
Hap	STRSELPKNS	GESNENWLYM	GKTSDEAKRN	VNMHINNERM	NGFNGYFGEE	
HK368IGA	SIRSEFPQNR	GESNNSWLYM	GTEKADAQKN	AMNHINNERM	NGFNGYFGEE	
HK393IGA	STRSELPKNS	GESNENWLYM	GKTSDEAKRN	VNMHINNERM	NGFNGYFGEE	
HK715IGA	STRSELPQNS	GESNENWLYM	GRTSDEAKRN	VNMHINNERM	NGFNGYFGEE	
HK61IGA	-----	-----	-----	-----	---	NG-FGE-	
Consensus							
701	D.KNKHNGRL	NLIYKPTTED	RTLLLSGGTN	LKGDITQTKG	KLFFSGRPTP	750
Hap	EGK..NNGNL	NVTFKGKSEQ	NRFLLTGGTN	LNGDLTVEKG	TLFLSGRPTP	
HK368IGA	EGK..NNGNL	NVTFKGKSEQ	NRFLLTGGTN	LNGDLNVQQG	TLFLSGRPTP	
HK393IGA	EGK..NNGNL	NVTFKGKSEQ	NRFLLTGGTN	LNGDLKVEKG	TLFLSGRPTP	
HK715IGA	ETKATQNGKL	NVTFNGKSDQ	NRFLLTGGTN	LNGDLNVEKG	TLFLSGRPTP	
HK61IGA	--K---NG-L	N-----	---LL-GGTN	L-GD-----	---	LF-SGRPTP	
Consensus							
751	HAYNHLNKRW	SEMEG..IPQ	GEIVWDHDWI	NRTFKAENFQ	IKGGSVVVS.	800
Hap	HARDIAGISS	TKKDPHFAEN	NEVVVEDDWI	NRNFKATTMN	VTGNASLYSG	
HK368IGA	HARDIAGISS	TKKDSHFSEN	NEVVVEDDWI	NRNFKATTMN	VTNNATLYSG	
HK393IGA	HARDIAGISS	TKKDQHFSEN	NEVVVEDDWI	NFNERATTMN	VTNNATLYSG	
HK715IGA	HARDIAGISS	TKKDPHFTEN	NEVVVEDDWI	NRNFKATTMN	VTGNASLYSG	
HK61IGA	HA-----	-----	-E-V---DWI	NR-FKA----	---	-----S-	
Consensus							

FIG._7D

Hap	801	RNVSSIEGNW	TVSNANATF	GVVPNQNTFI	CTRSDWTLGT	TCQKVDLTDT	850
HK368IGA		RNVANITSNI	TASNKAQVHI	GY..KTGDTV	CVRSDYTGIV	TCTTDKLS.	
HK393IGA		RNVESITSNI	TASNNAKVHI	GY..KAGDTV	CVRSDYTGIV	TCTTDKLS.	
HK715IGA		RNVANITSNI	TASDNAAVHI	GY..KAGDTV	CVRSDYTGIV	TCTTDKLS.	
HK61IGA		RNVANITSNI	TASNNAAVHI	GY..KTGDTV	CVRSDYTGIV	TCHNSNLSE.	
Consensus		RNV--I--N-	T-S--A----	G-----T-	C-RSD-TG--	TC-----L----	
				*	*	*	
Hap	851	KVINSIPKTQ	INGSINLTDN	ATANVKGLAK	LNGNVTLTNH	SQFTLSNNAT	900
HK368IGA		KALNSFNPTN	LRGNVNLTES	A.....	
HK393IGA		KALNSFNPTN	LRGNVNLTES	A.....	
HK715IGA		KALNSFNATN	VSGNVNLSGN	A.....	
HK61IGA		KALNSFNPTN	LRGNVNLTEN	A.....	
Consensus		K--NS---T-	--G--NL---	A-----	-----	-----	
Hap	901	QIGNIRLSDN	STATVDNANL	NGNVHLTDSA	QFSLKNSHFS	HQIQGDKGTT	950
HK368IGA	NFVLGKANL	FGTIQSRGNS	QVRLT.....	
HK393IGA	NFVLGKANL	FGTIQSRGNS	QVRLT.....	
HK715IGA	NFVLGKANL	FGTISGTGNS	QVRLT.....	
HK61IGA	SFTLGKANL	FGTIQSIGTS	QVNLK.....	
Consensus		-----	-----ANL	-G-----	Q--L-----	-----	
Hap	951	VTLENATWTM	PSDTTLQNL	LNNSTITLNS	AYSASSNNTF	RRRSLETETT	1000
HK368IGA		...ENSHWHL	TGNSDVHQLD	LANGHIHLNS	ADNSNNVTK.	
HK393IGA		...ENSHWHL	TGNSDVHQLD	LANGHIHLNS	ADNSNNVTK.	
HK715IGA		...ENSHMHL	TGDSNVNQLN	LDKGHIHLNA	QNDANKVTT.	
HK61IGA		...ENSHWHL	TGNSNVNQLN	LTNGHIHLNA	QNDANKVTT.	
Consensus		---EN--W--	-----L-	L-----I-LN-	-----	-----	

FIG..7E

Hap	1001	PTSAEHRFNT	LTVNGKLSGQ	GTFQFTSSLF	GYKSDKLKL	NDAEGDYILS	1050
HK368IGA	YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVT	KSATGNFTLQ	
HK393IGA	YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVT	KSATGNFTLQ	
HK715IGA	YNT	LTVNS.LSGN	GSFYLLTDLS	NKQGDKVVT	KSATGNFTLQ	
HK61IGA	YNT	LTVNS.LSGN	GSFYVVDFT	NNKSNKVVN	KSATGNFTLQ	
Consensus		-----NT	LTVN--LSG-	G-F-----	-----K----	--A-G---L-	
Hap	1051	VRNTGKEPET	LEQLTLVESK	DNQPLSDKLK	FTLENDHVDA	GALRYKLVKN	1100
HK368IGA		VADKTGEPNH	.NELTLFDAS	KAQR..DHLN	VSLVGNTVDL	GAWKYKLRNV	
HK393IGA		VADKTGEPNH	.NELTLFDAS	KAQR..DHLN	VSLVGNTVDL	GAWKYKLRNV	
HK715IGA		VADKTGEPNK	.NELTLFDAS	NATR..NNLN	VSLVGNTVDL	GAWKYKLRNV	
HK61IGA		VADKTGEPNH	.NELTLFDAS	NATR..NNLE	VTLANGSVDR	GAWKYKLRNV	
Consensus		V-----EP--	---LTL-----	-----L-	--L-----VD-	GA---YKL---	
Hap	1101	DGEFRLHNPI	KEQELHNDLV	1150
HK368IGA		NGRYDLYNP.	.EVEKRNQTV	DTTNITTPNN	IQADVPSVPS	NNEEIARVDE	
HK393IGA		NGRYDLYNP.	.EVEKRNQTV	DTTNITTPNN	IQADVPSVPS	NNEEIARVDE	
HK715IGA		NGRYDLYNP.	.EVEKRNQTV	DTTNITTPNN	IQADVPSVPS	NNEEIARV.E	
HK61IGA		NGRYDLYNP.	.EVEKRNQTV	DTTNITTPND	IQADAPSAQS	NNEEIARV.E	
Consensus		-G---L-NP-	-E-E--N--V	-----	-----	-----	
Hap	1151	1200
HK368IGA		APVPPAPAT	
HK393IGA		APVPPAPAT	
HK715IGA		TPVPPAPAT	
HK61IGA		TPVPPAPAT	
Consensus		-----	ESAIASEQPE	TRPAETAQPA	MEETNTANST	ETAPKSDTAT	

FIG._7F

Hap	1201	RAEQAERTLE	AKQVEPT	1250
HK368IGA	PSETTETVAE	NSKQESKTVE	KNEQDATETT	AQNREVAKEA
HK393IGA	PSETTETVAE	NSKQESKTVE	KNEQDATETT	AQNREVAKEA
HK715IGA	PSETTETVAE	NSKQESKTVE	KNEQDATETT	AQNGEVAEEA
HK61IGA	QENPNSESV	PSETTEKVAE	NPPQENETVA	KNEQEATEPT	PONGEVAKED
Consensus	-----	-----	-----	---Q---T---	-----T---	-----	-----
Hap	1251	AKTQT GE	1300
HK368IGA	KSNVKANTQT	NEVAQSGSET	KETQTTETK.	ETATVE
HK393IGA	KSNVKANTQT	NEVAQSGSET	KETQTTETK.	ETATVE
HK715IGA	KPNVKANTQT	NEVAQSGSET	EETQTTEIK.	ETAKVE
HK61IGA	OPTVEANTQT	NEATQSEGKT	EETQTAEIKS	EPTESVTVSE	NOPEKTVSQS
Consensus	-----	-----	-----	-----	-----	-----	-----
Hap	1301	1350
HK368IGA	KEEK.....
HK393IGA	KEEK.....
HK715IGA	KEEKAKVEKE	EKAKVEKDEI	QEAPQMASET	SPKQAKPAPK	EVSTDTKVEE
HK61IGA	TEDKVVVEKE	EKAKVETEET	QKAPQVTSKE	PPKQAEPAPE	EVPTDTNAEE
Consensus	-----	-----	-----	-----	-----	-----	-----
Hap	1351	1400
HK368IGA
HK393IGA
HK715IGA	TQVQAQPQTQ	STTVAAAEAT	SPNSKPAEET	.QPSEKTNAE	PVTPVSKNQ
HK61IGA	A..QALQQTQ	PTTVAAAEET	SPNSKPAEET	QQPSEKTNAE	PVTPVVS...
Consensus	-----	-----	-----	-----	-----	-----	-----

FIG._7G

1401	1450
HapPKVRS RRAARAAPFD TLP.....
HK368IGAAKVETE KTQEVPKVTS QVSPKQEQSE T.....
HK393IGAAKVETE KTQEVPKVTS QVSPKQEQSE T.....
HK71SIGA	TENTTDQPTREKTAKEVETE KTQEPPOVAS QASPKQEQSE T.....
HK61IGA	.ENTATQPTETEETAKVEKE KTQEVPOVAS QESPKQEQPA AKPQAQTKPQ
Consensus	-----P-V-S-----
1451	1500
Hap
HK368IGA
HK393IGAV
HK71SIGAV
HK61IGAV
Consensus	-----P-TGETAANS KPAAKPQAQA-----
1501	1550
HapD QSLNLEA.....KQAE L TAETQKSKAK TKK.....
HK368IGA	QPAEPAREN DPTVNIKEP.....QSQTNT TADTEQPAKE TSSNVE.....
HK393IGA	QPAEPAREN DPTVNIKEP.....QSQTNT TADTEQPAKE TSSNVE.....
HK71SIGA	QPAVLESEN VPTVNNAEV QAQLQTQTSATVSTKQPAPE NSINTG.....
HK61IGA	KPQTEPAREN VSTVNTKEP.....QSQTSA TVSTEQPAKE TSSNVEQPAP
Consensus	-----N-E-----Q-T-----
1551	1600
HapV RSKRAVFSDP LLDQSL.....
HK368IGAQPVT ESTTVNTGNS VVEN.....
HK393IGAQPVT ESTTVNTGNS VVEN.....
HK71SIGASAT AITETAEKSD KPQTETAAST EDASQHKANT VADNSVANNNS
HK61IGA	ENSINTGSAT TMTETAEKSD KPQMET..VT ENDRQPEANT VADNSVANNNS
Consensus	-----

FIG._7H

1601

 ESSEPKSRRR RSISQPQETS
 ESSEKSRRR RSVSQPKETS

 1650
 ALEAALEVID APQQSEKDRL AQEEAEKQRK
 PENTTPATTQ PTVNSESSN. .KPK.NRHR
 PENTTPATTQ PTVNSESSN. .KPK.NRHR
 AEETTAASTD ETTIADNSKR SKPN.RRSRR
 AEETTVASTQ ETTVDNSVST PKPRSRRTRR
 -----R-

1651

 SVRSVPHNVE PATSSND..
 SVRSVPHNVE PATSSND..
 SVRS.....E PTVTNGSD..
 SVQTNSYEPV ELPTENAENA ENVQSGNNVA
 -----L-
 1700
 SRYNSALSE
 STNTNAVLSD
 STNTNAVLSD
 STNTNAVLSD
 STNTNAVLSD
 STNTNAVLSD
 SKNTNAVLSD
 S-N-S-

1701
 LSA.....TV NSMLSVQDEL DRL.FVDQAAQ SAVWTNIAQD KRRYDSDAFR
 ARAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWVSNTSM NKNYSSSQYR
 ARAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWVSNTSM NKNYSSSQYR
 AMAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWVSNTSM NENYSSSQYR
 AMAKAQFVAL NVGKAVSQHI SQLEMNNEGQ YNVWISNTSM NKNYSSSEQYR
 -----A-----N-----V-----L-----Q-----Y-S-----R

1751
 AYQQQKTNLR QIGVQKALAN GRIGAVFSHS RSDNTFDEQV KNHATLTMS
 RFSSKSTQTQ LGWDQTISNN VOLGGVFTYV RNSNFDKAT SKN.TLAQVN
 RFSSKSTQTQ LGWDQTISNN VOLGGVFTYV RNSNFDKAT SKN.TLAQVN
 RFSSKSTQTQ LGWDQTISNN VOLGGVFTYV RNSNFDKAS SKN.TLAQVN
 RFSSKSTQTQ LGWDQTISNN VOLGGVFTYV RNSNFDKAS SKN.TLAQVN
 -----T-----Q-----N-----G-VF-----R-N-FD-----TL-----

FIG. 71

Hap	1801	GFAQYQWGD	QF..GVNVGT	GISASKMAEE	QSRKIHRKAI	NYGVNASYQF	1850
HK368IGA		FYSKY.YADN	HWYLGIDLG	GKFQSKLQTN	HNAKFARHTA	QFGLTAGKAF	
HK393IGA		FYSKY.YADN	HWYLGIDLG	GKFQSKLQTN	HNAKFARHTA	QFGLTAGKAF	
HK715IGA		FYSKY.YADN	HWYLGIDLG	GKFQSNLKTN	HNAKFARHTA	QFGLTAGKAF	
HK61IGA		FYSKY.YADN	HWYLGIDLG	GKFQSNLQTN	NNAKFARHTA	QIGLTAGKAF	
Consensus		----Y---D-	----G---G-	G---S-----	---K--R---	--G--A---F	
Hap	1851	RLGQLGIQPY	FGVNRYFIER	ENYQSEEV RV	KTPSLAFNRY	NAGIRVDYTF	1900
HK368IGA		NLGNFGITPI	VGVRYSYLSN	ADFALDQARI	KVNPISVKTA	FAQVDLSYTY	
HK393IGA		NLGNFGITPI	VGVRYSYLSN	ADFALDQARI	KVNPISVKTA	FAQVDLSYTY	
HK715IGA		NLGNFGITPI	VGVRYSYLSN	ANFALAKDRI	KVNPISVKTA	FAQVDLSYTY	
HK61IGA		NLGNFAVKPT	VGVRYSYLSN	ADFALAQDRI	KVNPISVKTA	FAQVDLSYTY	
Consensus		-LG-----P-	-GV-----	-----R-	K-----	-A-----YT-	
Hap	1901	TPTDNISVKP	YFFVNYVDVS	NANVQTTVNL	TVLQQPFGRY	WQKEVGLKAE	1950
HK368IGA		.HLGEFSVTP	ILSARY.DAN	QSGGKINVG	YDFAYNVENQ	QQYNAGLKLK	
HK393IGA		.HLGEFSVTP	ILSARY.DAN	QSGGKINVG	YDFAYNVENQ	QQYNAGLKLK	
HK715IGA		.HLGEFSVTP	ILSARY.DTN	QSGGKINVG	YDFAYNVENQ	QQYNAGLKLK	
HK61IGA		.HLGEFSITP	ILSARY.DAN	QNGGKINVS	YDFAYNVENQ	QQYNAGLKLK	
Consensus		-----S--P	-----Y-D--	-----V--	-----	-Q---GLK--	
Hap	1951	ILHFQISAFI	SKSQGSQGLK	QQNVGVKLG	Y		1982
HK368IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	RW		
HK393IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
HK715IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
HK61IGA		YHNVKLSLIG	GLTKAKQAEK	QKTAELKLSF	SF		
Consensus		-----S---	-----Q--K	Q-----KL--	--		

FIG..7J

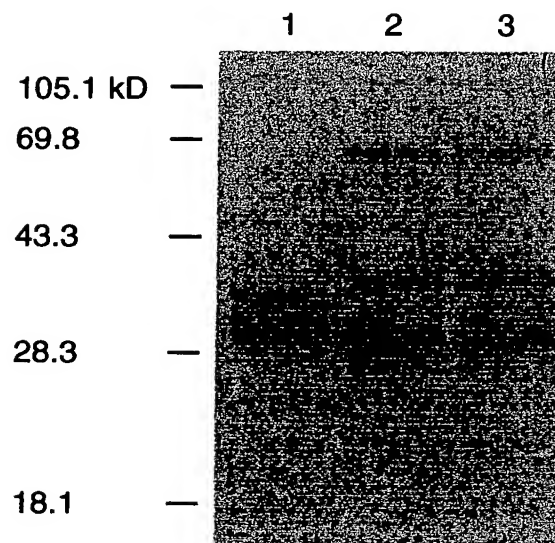


FIG._8

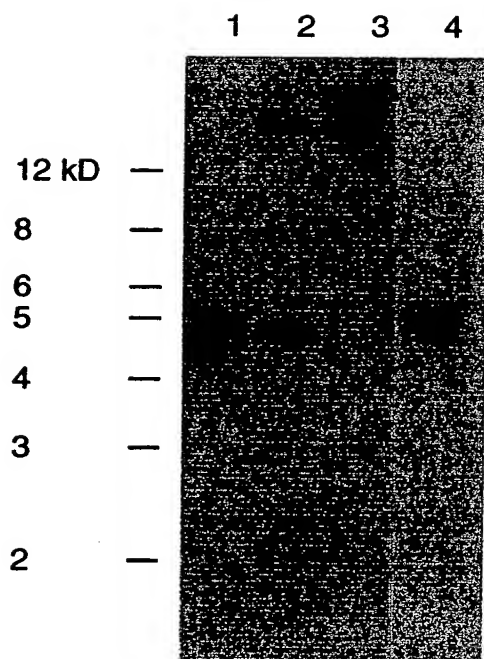


FIG._9A

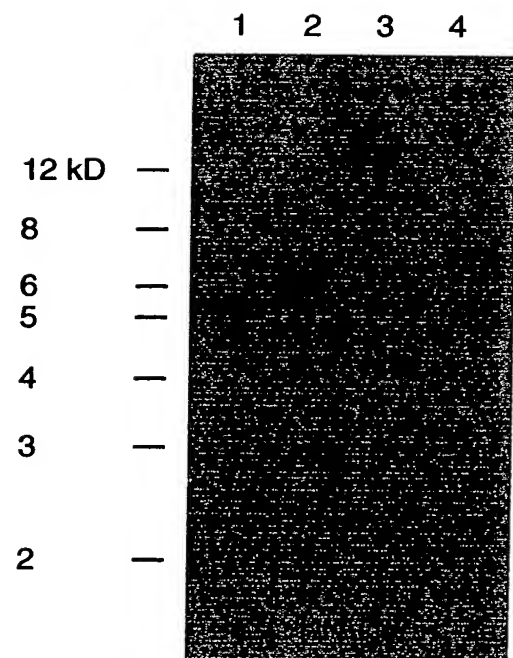


FIG._9B

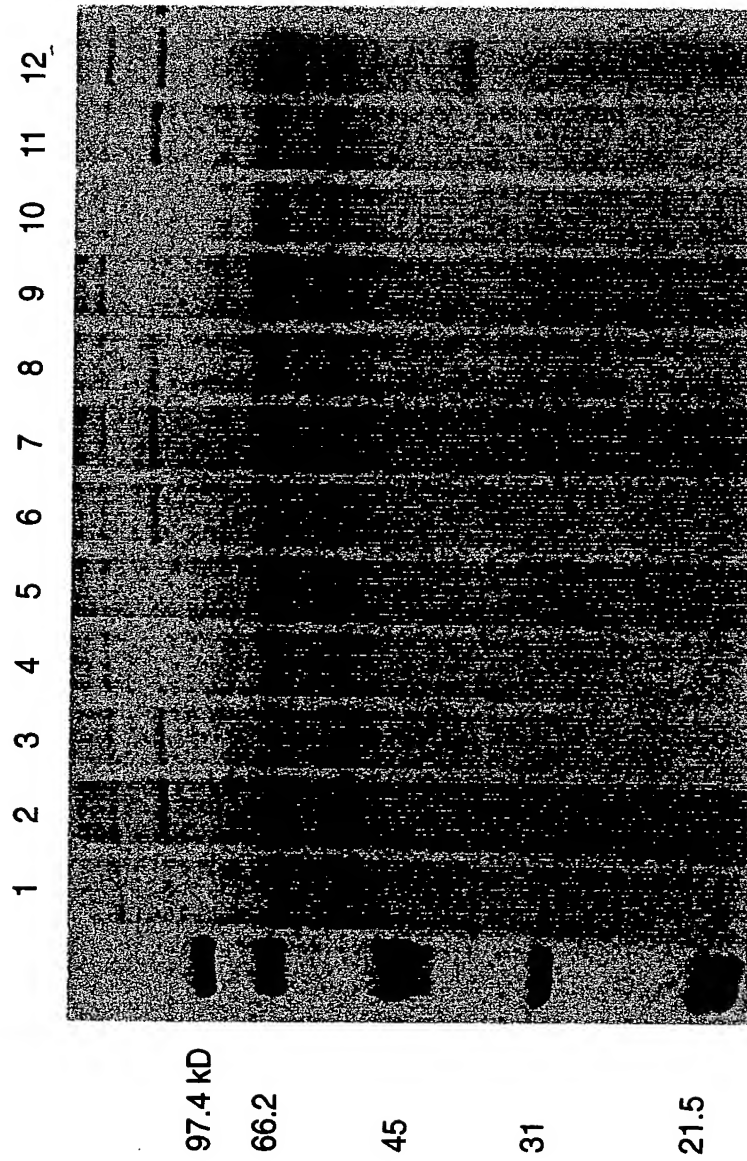


FIG. 10

24 / 45

		1		50
HapN187	(1)	MKKT	VFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYR	DF
HapTN106	(1)	MKKT	VFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYR	DF
Hap860295	(1)	MKKT	VFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYR	DF
Consensus	(1)	MKKT	VFRLNFLTACISLGIVSQAWAGHTYFGIDYQYYR	DF
		51		100
HapN187	(51)	AQNI	KVYNKQGLVGTSM	TKAPMIDFSV
HapTN106	(51)	AQDI	IYNKKGEMIGTMM	KGVPMFDLSS
Hap860295	(51)	AKNI	EVYNKEGTLVGT	SMTKAPMIDFS
Consensus	(51)	A I	YNK G GT M	PM D S R G L Q
		101		150
HapN187	(101)	GYTD	VDFGAEGNPNPDQ	HRFTYKIVKRNNY
HapTN106	(101)	GYDV	DFGMEGENPDQHR	FYKIVKRYNYKSG
Hap860295	(101)	GYNS	VDFGAEGNPNPDQ	HRFTYQIVKRNNY
Consensus	(101)	GY	VDFG EG NPDQ	HRF Y VKR NYK
		151		200
HapN187	(151)	TEAAP	IDMTSNMNGSTYS	DRTKYPERVRIGS
HapTN106	(141)	TETAP	LEMVSYMDCNHYK	NFNQYPLRVRVGS
Hap860295	(151)	TDAEP	AKMTDNMNGKNY	ADLSKYFDRVRIG
Consensus	(151)	T	P M M G Y YP	RVR G G Q W D G
		201		250
HapN187	(196)	VAGAY	HYLTAGNTHNORG	AGNGYSYLGGDV
HapTN106	(195)	LAYGG	SWLIGGNTFEDG	PAGNGTLELNGR
Hap860295	(201)	LADAY	LWRIAGNTHSOS	GAGNGTVNLSGD
Consensus	(201)	A	GNT AGNG	L G YGPLP S
		251		300
HapN187	(246)	PMFI	YDAEKQKWLING	LLREGNPFEGKENG
HapTN106	(245)	PMFI	YDKKWKWLLNG	VLRGNHYAAVGN
Hap860295	(251)	PMFI	YDAIKQKWLING	VLOTGNPFSGAG
Consensus	(251)	PMFI	YD KWL NG L	GNP N Q RK F D
		301		350
HapN187	(295)	SLYTR	AGNGVYTHSQND	NG-----QGS
HapTN106	(294)	NFWDT	NAEYRFNIGSDH	NGRVATIKSTLP
Hap860295	(301)	TFL	EPSRNGHYSFTS	NMNG-----TGT
Consensus	(301)			NG L
		351		400
HapN187	(340)	-EKK	KVHNPRYDGE-	NTYSPRLNNGET
HapTN106	(344)	RDKN	GDESPSYKGE-	NPWSPALHHGKS
Hap860295	(346)	ALKE	KDKEPVYAAGGV	NAYKPRLLNNGKN
Consensus	(351)	K	P Y N P L G	YF D G L INQAG

FIG. 11A

25 / 45

		401		450
HapN187	(387)	GLYFEGNFTVSPNSN-QTWQGAGLHVSENSTVTWKVNGVEHDRLSKIGKG		
HapTN106	(392)	GLYFEGNFVVKGNQNNITWQGAGVSVGEESTVEWQVHNPEGDRLSKIGLG		
Hap860295	(396)	GLYFEGNFTVSSENN-ATWQGAGVHVGEDSTVTWKVNGVEHDRLSKIGKG		
Consensus	(401)	GLYFEGNF V N TWQGAG V E STV W V E DRLSKIG G		
		451		500
HapN187	(436)	TLHVQAKGENKGSISVGDGKVILEQQADDQGNKQAFSEIQLVSGRGTVQL		
HapTN106	(442)	TLLVNGKGNLGLSVGNGLVLDQQADESGQKQAFKEVGIVSGRATVQL		
Hap860295	(445)	TLHVQAKGENLGSISVVGKGVILDQQADENNQKQAFKEVGIVSGRATVQL		
Consensus	(451)	TL KG N GS SVG G V L QQAD KQAF E G VSGR TVQL		
		501		550
HapN187	(486)	NDKQFDITDKFYFGFRGGRDLNGHSLTFKRIQNTDEGAMIVNHNNTTQAA		
HapTN106	(492)	NSADQVDPNNIYFGFRGGRDLNGHSLTFERIQNTDEGAMIVNHNASQTA		
Hap860295	(495)	NSADQVDPNNIYFGFRGGRDLNGHSLTFKRIQNTDEGAMIVNHNNTTQVA		
Consensus	(501)	N Q D YFGFRGGRDLNGHSLTF RIQNTDEGAMIVNHN Q A		
		551		600
HapN187	(536)	NVTITGNESIVLP-NGNNINKLDYRKEIAYNGWFGETDKNKHNGRLNLIY		
HapTN106	(542)	NITITGNATLNS-----DSKQLTNKKDIAFNGWFGEOQDKAKTNGRLNVNY		
Hap860295	(545)	NITITGNESITAPSNKNNINKLDYSKEIAYNGWFGETDENKHNGRLNLIY		
Consensus	(551)	N TITGN I L K IA NGWFG E D K NGRLN Y		
		601		650
HapN187	(585)	KPTTEDRTL LLLSGGTNLKGDITQTKGKLFFSGRPTPHAYNHLNKRWSME		
HapTN106	(587)	QPVNAENH LLLSGGTNLNGNITONGGTLVFSGRPTPHAYNHLRDLNME		
Hap860295	(595)	KPTTEDRTL LLLSGGTNLKGNITQEGGTLVFSGRPTPHAYNHLNR--PNEL		
Consensus	(601)	P LLLSGGTNL G ITQ G L FSGRPTPHAYNHL		
		651		700
HapN187	(635)	GIPQGEIVWDHWDWINRTFKAENFQIKGGS AVVSRNVSSIEGNWTVSNNAN		
HapTN106	(637)	GIPQGEIVWDHWDWINRTFKAENFQIKGGS AVVSRNVSSIEGNWTVSNNAN		
Hap860295	(643)	GRPOGEIVWDHWDWINRTFKAENFQIKGGS AVVSRNVSSIEGNWTVSNNAN		
Consensus	(651)	G PQGE V D DWI RTFKAENFQIKGGS AVVSRNVSSIEGNWTVSNNAN		
		701		750
HapN187	(685)	ATFGVVPNQONTICTRSDWTGLTTCQKVDLTDTKVINSIPTQTQINGSINL		
HapTN106	(687)	ATFGVVPNQONTICTRSDWTGLTTCQKVDLTDTKVINSIPTQTQINGSINL		
Hap860295	(693)	AATFGVVPNQONTICTRSDWTGLTTCQKVDLTDTKVINSIPTQTQINGSINL		
Consensus	(701)	A FGVVPNQONTICTRSDWTGLTTC VDLTD KVINSIP TQINGSINL		
		751		800
HapN187	(735)	TDNATANVKG LAKLNGNVTLINHSQFTLSNNATQTGNIRLS DNSTATVDN		
HapTN106	(737)	TDNATVNIHGLAKLNGNVTLINHSQFTLSNNATQTGNIRLS DNSTATVDN		
Hap860295	(743)	TDNATVNIHGLAKLNGVVTLINHSQFTLSNNATQTGNIRLS DNSTATVDN		
Consensus	(751)	TDNAT N GLAKLNGNVTL HSQFTLSNNATQ GNI LS ATVDN		

FIG. 11B

		801		850
HapN187	(785)	ANLNGNVHLTDSAQFSLKNSHFSHQIQGDKGTTVTLENATWTMPSDTTLQ		
HapTN106	(787)	ANLNGNVHLMDSAQFSLKNSHFSHQIQGGEDTTVMLENATWTMPSDTTLQ		
Hap860295	(793)	ANLNGNVHLTDSAQFSLKNSHFSHQIQGDKDTTVTLENATWTMPSDATLQ		
Consensus	(801)	ANLNGNV L DSAQFSLKNSHFSHQIQG TTV LENATWTMPSD TLQ		
		851		900
HapN187	(835)	NLTLNNSTITLNSAYSASSNNTPRRRRRSLETETTP TSAEHRFNTLT VNG		
HapTN106	(837)	NLTLNNSTYTLNSAYSASISNNAPRRRRRSLETETTP TSAEHRFNTLT VNG		
Hap860295	(843)	NLTLNNSTYTLNSAYSASSNNAPR-HRRSLETETTP TSAEHRFNTLT VNG		
Consensus	(851)	NLTLNNST TLNSAYSAS SNN PR RRSLETETTP TSAEHRFNTLT VNG		
		901		950
HapN187	(885)	KLSGQGTFTQFTSSSLFGYKSDKLKLSNDAEGDYLLSVRNTGKEPETLEQLT		
HapTN106	(887)	KLSGQGTFTQFTSSSLFGYKSDKLKLSNDAEGDYTL SVRNTGKEPETLEQLT		
Hap860295	(892)	KLSGQGTFTQFTSSSLFGYKSDKLKLSNDAEGDYTL SVRNTGKEPETLEQLT		
Consensus	(901)	KLSGQGTFTQFTSSSLFGYKSDKLKLSNDAEGDY LSVRNTGKEP QLT		
		951		1000
HapN187	(935)	LVESKDNQPLSDKLKFTLENDHVDAGALRYKLVKNDGEFRLHNP I KEQEL		
HapTN106	(937)	LVESKDNKPLSDKLFTLENDHVDAGALRYKLVKNDGEFRLHNP I KEQEL		
Hap860295	(942)	LVESKDNKPLSDKLKFTLENDHVDAGALRYKLVKNNGEFRLHNP I KEQEL		
Consensus	(951)	LVESKDN PLSDKL FTLENDHVDAGALRYKLVKN GEFRLHNP I KEQEL		
		1001		1050
HapN187	(985)	HNDLVRAEQAERTLEAKQVEPTAKTQTGEPKVR SRRAARAFD TLPDQS		
HapTN106	(987)	RSDLVRAEQAERTLEAKQVEQTAKTQTSKARVRSR--RAVFS D TLPDQS		
Hap860295	(992)	RNDLVRAEQAERTLEAKQVEQTAEQTQTSNARVRSK--RAVFS D TLPDQS		
Consensus	(1001)	DLVRAEQAERTLEAKQVE TA TQT VRS RA F D LP QS		
		1051		1100
HapN187	(1035)	LLNALEAKQAELTAETQKSKAKTKKVR SKRAV--FSDPLLDQS-----		
HapTN106	(1034)	LLKALEAKQA-LTTETQTS--KAKKVR SKRAAREFSD TLPDQ-----		
Hap860295	(1039)	QLDVLQAEQVEPTAEKQKN--KAKKVR SKRAV--FSD TLPDQSOLDVLQA		
Consensus	(1051)	L L A Q T E Q K KVR SKRA FSD L DQ		
		1101		1150
HapN187	(1076)	-----LFALEAALEVIDAQQ		
HapTN106	(1073)	-----LQAAL EVIDAQQ		
Hap860295	(1085)	EQVEPTAEKQKNKAKKVR SKRAAREFSD TPLDLSRLKVL EKVLEVIN AQQ		
Consensus	(1101)	L LEVI A Q		
		1151		1200
HapN187	(1091)	QSEKDRLAQEEAEK-QRKQKDLISRYSNSALS ELSATVNSMLS VQDELDR		
HapTN106	(1086)	QVKKEPQIQEEEEKRQRKQKELISRYSNSALS ELSATVNSMLS VQDELDR		
Hap860295	(1135)	QVKKEPQDQ--EK-QRKQKDLISRYSNSALS ELSATVNSMLS VQDELDR		
Consensus	(1151)	Q K Q EK QRKQK LISRYSNSALS ELSATVNSMLS VQDELDR		

FIG. 11C

27 / 45

		1201		1250
HapN187	(1140)	LFVDQAQSAVWTNIAQDKRRYDSDAFRAYQQKTNLRQIGVQKALANGRIG		
HapTN106	(1136)	LFVDQAQSAVWTNIAQDKRRYDSDAFRAYQQKTNLRQIGVQKALANGRIG		
Hap860295	(1181)	LFVDQAQSAVWTNIAQDKRRYDSDAFRAYQQKTNLRQIGVQKALANGRIG		
Consensus	(1201)	LFVDQAQSAVWTNIAQDKRRYDSDAFRAYQQKTNLRQIGVQKAL NGRIG		
		1251		1300
HapN187	(1190)	AVFSHSRSDNTFDEQVKNHATLTMMSGFAQYQWGDLOFGVNVGTGISASK		
HapTN106	(1186)	AVFSHSRSDNTFDEQVKNHATLTMMSGFAQYQWGDLOFGVNVGTGISASK		
Hap860295	(1231)	AVFSHSRSDNTFDEQVKNHATLTMMSGFAQYQWGDLOFGVNVGTGISASK		
Consensus	(1251)	AVFSHSRSDNTFDEQVKNHATL MMSGFAQYQWGDLOFGVNVG GISASK		
		1301		1350
HapN187	(1240)	MAEEQSRKIHRRKAINYGVNASYQFRLGQLGIQPYFGVNRYFIERENYQSE		
HapTN106	(1236)	MAEEQSRKIHRRKAINYGVNASYQFRLGQLGIQPYFGVNRYFIERENYQSE		
Hap860295	(1281)	MAEEQSRKIHRRKAINYGVNASYQFRLGQLGIQPYFGVNRYFIERENYQSE		
Consensus	(1301)	MAEEQSRKIHRRKAINYGVNASYQFRLGQLGIQPY GVNRYFIERENYQSE		
		1351		1400
HapN187	(1290)	EVVKVTPSLAFNRYNAGIRVDYTFPTDNISVKPYFFVNYVDVSNANVQT		
HapTN106	(1286)	EVVKVTPSLAFNRYNAGIRVDYTFPTDNISVKPYFFVNYVDVSNANVQT		
Hap860295	(1331)	EVVKVTPSLAFNRYNAGIRVDYTFPTDNISVKPYFFVNYVDVSNANVQT		
Consensus	(1351)	EV V TPSL FNRYNAGIRVDYTFPTDNIS KPYFFVNYVDVSNANVQT		
		1401		1450
HapN187	(1340)	TVNLTVLQQPFGRYWQKEVGLKAEILHFQISAFISKSQGSQGLGKQQNVGV		
HapTN106	(1336)	TVNLTVLQQPFGRYWQKEVGLKAEILHFQISAFISKSQGSQGLGKQQNVGV		
Hap860295	(1381)	TVNLTVLQQPFGRYWQKEVGLKAEILHFQISAFISKSQGSQGLGKQQNVGV		
Consensus	(1401)	TVN T LQQ FGRYWQKEVGLKAEILHFQ SAFISKSQGSQGLGKQQNVGV		
		1451		
HapN187	(1390)	KLYGRW		
HapTN106	(1386)	KLYGRW		
Hap860295	(1431)	KLYGRW		
Consensus	(1451)	KLYGRW		

FIG._11D

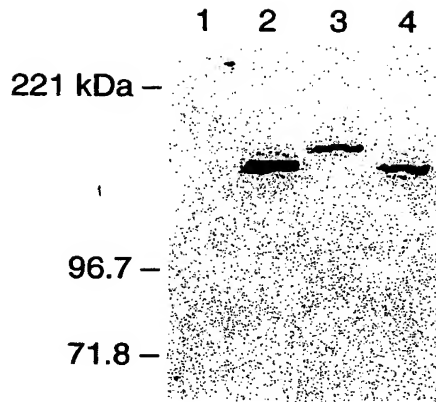


FIG._12A

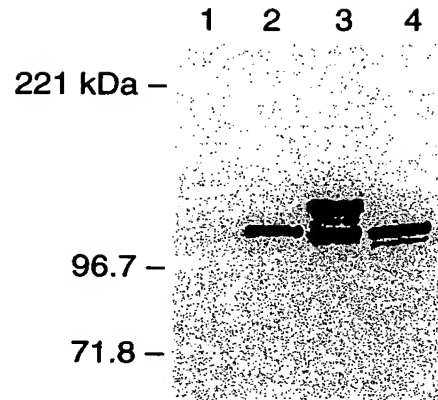


FIG._12B

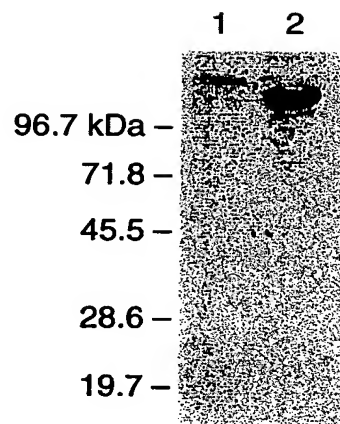
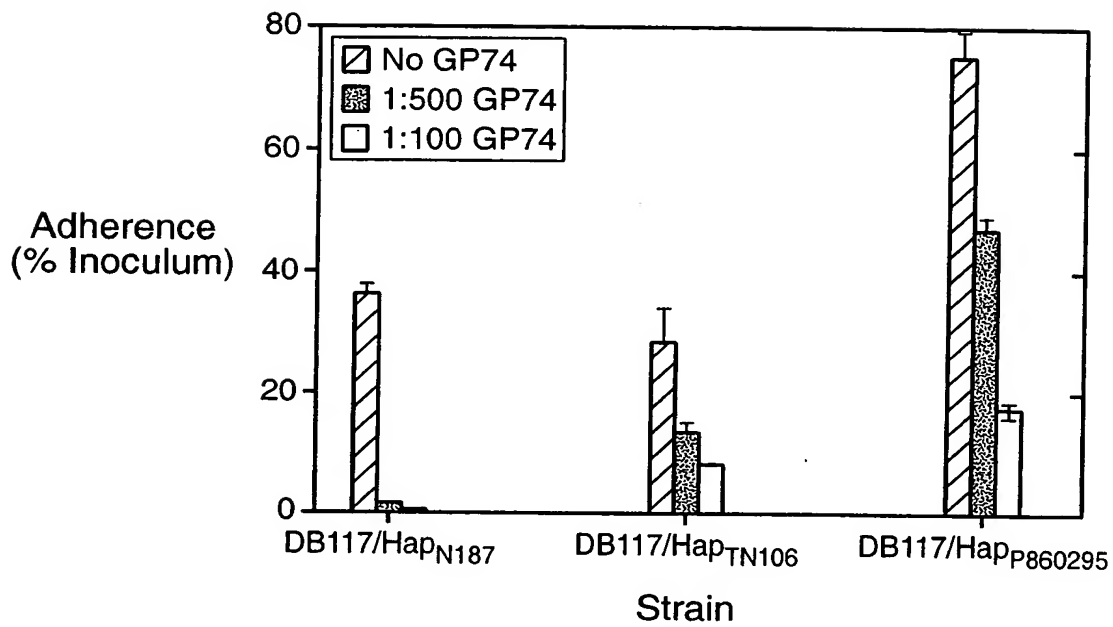
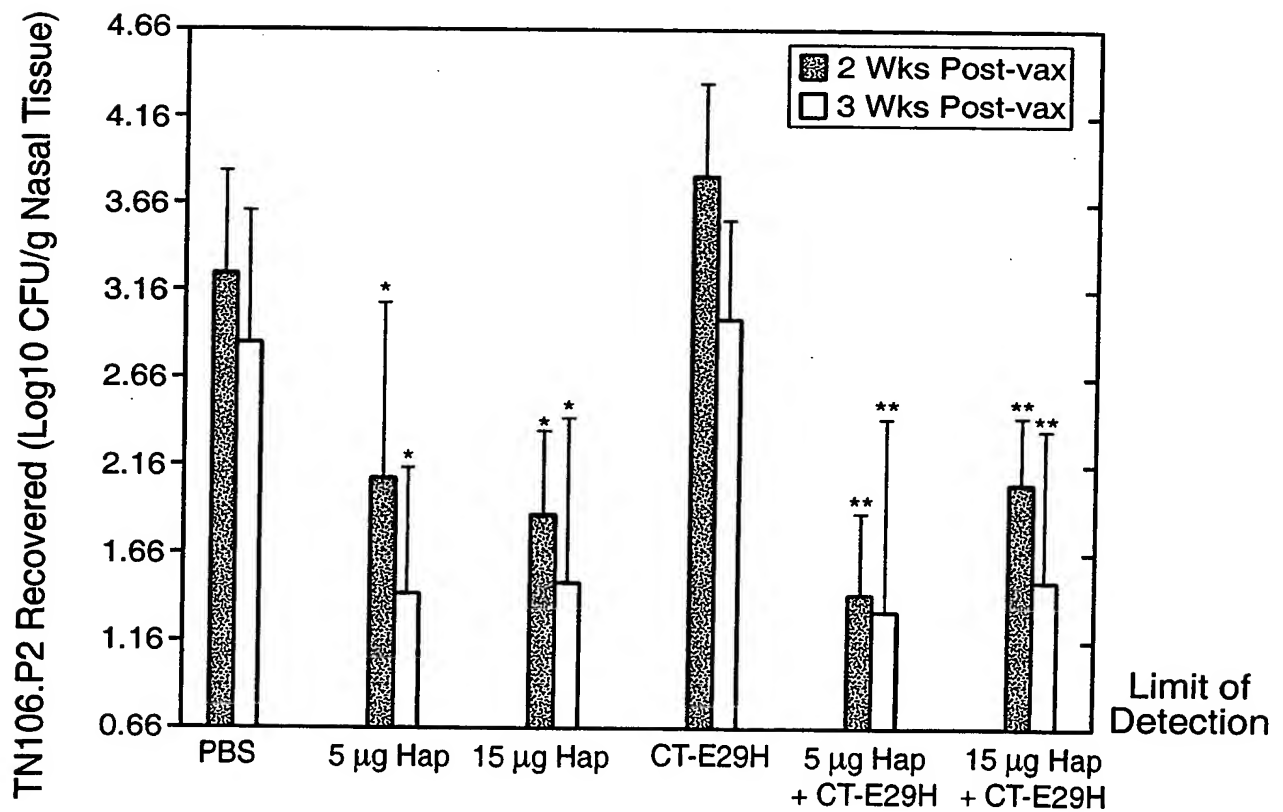


FIG._14

**FIG. 13****FIG. 15**

30 / 45

Nucleotide sequence for NTHi strain 11 hap gene (start codon to stop codon):

```

1  ATGAAAAAAA CTGTATTTTCG TCTTAATTTT TTAACCGCTT GCATTTTCATT
51  AGGGATAGTA TCGCAAGCGT GGGCAGGTCA TACTTATTTT GGGATTGACT
101 ACCAATATTA TCGTGATTTT GCCGAGAATG AAGGCAAGTT TGCAGTTGGG
151 GCTAAAAATA TTGATGTTTA TAACAAAGAA GGGCAATTAG TTGGCACATC
201 AATGACAAAA GCCCCGATGA TTGATTTCTC AGTCGTTTCC AGAAATGGAG
251 TTGCTGCCTT AGTAGGCGAT CAGTATATTG TGAGTGTGGC ACATAATGTA
301 GGCTATACCA ATGTGGATTT TGGTGCTGAA GGACAAAATC CTGATCAACA
351 TCGTTTTACT TATAAAATTG TGAAACGGAA TAATTATAAT CACGATGCGA
401 AGCACCGCTA TCTAGATGAC TACCATAATC CACGTTTACA TAAATTTGTA
451 ACGGATGCGG CACCAATTGA TATGACTTCA CATATGGATG GCAATAAGTA
501 TGCAAATAAG GAAAAATATC CTGAACGAGT ACGCGTCGGA TCTGGAGATC
551 AGTATTGGGA TGACGATCAA AACAAACAGAA CTTATTTATC TGACGGATAT
601 AATTATTTAA CAGGTGGGAA TACATATAAT CAAAGCGGTA GAGGTGATGG
651 ATATTCATAT GTGAGAGGTG ATATTCGCAA AGTTGGCGAT TATGGTCCAT
701 TACCGATTGC AAGTTCATTC GGGGACAGTG GATCTCCAAT GTTTATTTAT
751 GATGCTGAAA CACAAAAATG gCTAATTAAT GGAGTATTGC GGGAGGGGCA
801 ACCTTATACA GGCGAATTCG ATGGATTTCA ATTAGCCCGT AAATCTTTCC
851 TTGATGAAAT TATACGCAA GATCAACCAA ATGGTTTTTT AACCCTAAG
901 GGAATGGCG TTTATACCAT TTCTAAAAGT GACGATGGGA TAGGAGTTGT
951 TACTTCGAAA ATTGGAAAAC CTCGTGAAAT ACCTTTAGCG AACAACAAAT
1001 TAAAAATAGA AGATAAAGAT ACTGTCTATA ATAACAGATA TAATGGTCCT
1051 AATATTTATT CTCCTCAATT AAACAATGGC AAGAATATTT ATTTTGGAGA
1101 TGAAGAATTA GGATCCATAA CTTTAACGAC TGATATCGAT CAAGGTGCAG
1151 GCGGTTTGTA TTTTGAGGGG GATTTTATAG TTTCGCCTAC CAAAAATGAA
1201 ACGTGGAAG GTGCGGGCAT TCATGTCAGT GAAATTAGTA CCGTTACTTG
1251 GAAAGTAAAC GGCGTGGAAT ATGATCGACT TTCTAAAATC GGTAAAGGAA
1301 CATTACACGT TAAAGCCAAA GGGGAAAATA AAGGTTTCAT CAGCGTAGGC
1351 GATGGTAAAG TCATTTTGGA GCAGCAGGCA GACGATCAAG GCAACAAACA
1401 AGCCTTTAGT GAAATTGGCT TGGTTAGCGG CAGAGGGACT GTTCAATTAA
1451 ACGATGATAA ACAATTTGAT ACCGATAAAT TTTATTTTCG CTTTCGTGGT
1501 GGTGCTTAG ATCTTAACGG ACATTCATTA ACCTTTAAAC GTATCCAAAA
1551 TACGGACGAG GGGGCGATGA TTGTGAACCA TAATACAAC TCAAGTCGCTA
1601 ATATTACTAT TACTGGGAAC GAAAGTATTA CTGCTCCATC TAATAAAAAT
1651 AATATTAATA AACTTGATTA CAGCAAAGAA ATTGCCTACA ACGGCTGGTT
1701 TNGCGAAACA GATAAAAATA AACATAATGG ACGATTAAAC CTTATTTATA
1751 AACCAACCAC AGAAGATCGT ACTTTGCTAC TTTCAGGCGG CACAACTTA
1801 AAAGGCGATA TACTCAAAC AAAAGGTAAA CTATTTTTC TCGGTTAGACC
1851 GACACCCAC GCCTACAATC ATTTAGACAA ACGTTGGTCA GAAATGGAAG
1901 GTATCCCACA AGGCGAAATT GTGTGGGATT ACGATTGGAT TAACCGCACA
1951 TTTAAAGCTG AAAACTTCCA AATTAAAGGC GGAAGTGCGG TGGTTTCTCG
2001 CAATGTTTCT TCAATTGAGG GAAATTGGAC AGTCAGCAAT AATGCAAATG

```

FIG. 16A

2051 CCACATTTGG TGTGTGCCA AATCAGCAAA ATACCATTTG CACGCGTTCA
 2101 GATTGGACAG GATTAACGAC TTGTAACAACA GTTAATTTAA CCGATAAAAA
 2151 AGTTATTGAT TCCATACCGA CAACACAAAT TAATGGTTCT ATTAATTTAA
 2201 CTGATAATGC AACAGTGAAT ATTAATGGTT TAGCAAAACT TAATGGTAAT
 2251 GTCACTTTAA TAAATCATAG CCAATTTACA TTGAGCAACA ATGCCACCCA
 2301 AATAGGCAAT ATCAAACCTT CAAATCACGC AAATGCAAGG GTAAATAATG
 2351 CCACTTTAAT GGGCGATGTG AATTTAGCGG ATACTAGCCG TTTTACATTA
 2401 AGCAATCAAG CAACACAGAT TGGCACAATC AGTCTTCATC AGCAAGCTCA
 2451 AGCAACAGTG GATAATGCAA ACTTGAACGG TAATGTGCAT TTAACGGATT
 2501 CTGCCAGATT TTCTTTAAAA AACAGTCATT TTTCGCACCA AATTCAGGGC
 2551 GACAAAGACA CAACAGTGAC GTTGGAAAAT GCGACTTGGA CAATGCCTAG
 2601 CGATACTACA TTGCAGAATT TAACGCTAAA TAATAGTACT GTTACGTTAA
 2651 ATTCAGCTTA TTCAGCTAGC TCAAATAATG CGCCACGTCG CCgCCGTTCA
 2701 TTAGAGACGG AAACAACGCC AACATCGGCA GAACATCGTT TCAACACATT
 2751 GACAGTAAAT GGTAAATTGA GCGGGCAAGG CACATTCCAA TTTACTCCAT
 2801 CTTTATTTGG CTATGAAAGC GATAAATTAA AATTATCCAA TGACGCTGAG
 2851 GGCGATTACA CATTATCTGT TCGCAACACA GGCAAAGAAC CCGTGACCCT
 2901 TGAGCAATTA ACTTTGGTTG AAAGCAAAGA TAATAAACCG TTATCAGACA
 2951 AACTCAAATT TACTTTAGAA AATGACCACG TTGATGCAGG TGCATTACGT
 3001 TATAAATTAG TGAAGAATAA GGGCGAATTC CGCTTGCATA ACCCAATAAA
 3051 AGAGCAGGAA TTGCGCTCTG ATTTAGTAAG AGCAGAGCAA GCAGAACGAA
 3101 CATTAGAAGC CAAACAAGTT GAACAGACTG CTGAAACACA AACAAAGTAAT
 3151 GCAAGAGTGC GGTCAAGAAG AGCGGTGTTG TCTGATACCC CGTCTGCTCA
 3201 AAGCCTGTTA AACGCATTAG AAGTCAAACA AGCTGAACCG AATGCTAAAA
 3251 CACAAAAAAG TAAGGCAAAA ACAAAAAAAG CGCGGTCAAA AAGAGCATTG
 3301 AGAGAAGCGT TTTCTGATAC CCCGCCTGAT CTAAGCCAGT TAAACGTATT
 3351 AGAAGCCGCA CTTAAGGTTA TTAATGCCCA ACCGCAAACA GAAAAAGAAC
 3401 GTCAAGCTCA AGAGGAAGAA GCGAAAAGAC AACGCaACA AAAAGACTTG
 3451 ATCAGCCGTT ACTCAAATAG TGC GTTATCG GAGTTGTCTG CAACAGTAAA
 3501 TAGTATGCTT TCCGTTCAAG ATGAATTGGA TCGTCTTTTT GTAGATCAAG
 3551 CACAATCTGC CCTGTGGACA AATATCGCAC AGGATAAAAG ACGCTATGAT
 3601 TCTGATGCGT TCCGTGCTTA TCAGCAGAAA ACGAACTTGC GTCAAATTGG
 3651 GGTGCAAAAA GCCTTAGATA ATGGACGAAT TGGGGCGGTT TTCTCGCATA
 3701 GCCGTTTACA TAATACCTTT GACGAACAGG TTAAAAATCA CGCGACATTA
 3751 ACGATGATGT CGGGTTTTGC CCAATATCAA TGGGGCGATT TACAATTGGG
 3801 TGTAACCGTG GGCGCGGGAA TTAGTGCGAG TAAATGGCT GAAGAACAAA
 3851 GCCGAAAAAT TCATCGAAAA GCGATAAATT ATGGTGTGAA TGCAAGTTAT
 3901 CAGTTCCGTT TAGGGCAATT GGGTATTCAG CTTATTTGG GTGTAAATCG
 3951 ATATTTTATT GAACGTGAAA ATTATCAATC TGAAGAAGTG AAAGTGCAAA
 4001 CACCGAGCCT TGCATTTAAT CGCTATAATG CTGGCATTCT AGTTGATTAT
 4051 ACATTTACCC CGACAGATAA TATCAGCGTT AAGCCTTATT TCTTTGTCAA
 4101 TTATGTTGAT GTTTCAAACG CTAACGTACA AACCCTGTA AATAGCACGA
 4151 TGTTGCAACA ATCATTGTTGG CGTTATTGGC AAAAAGAAGT GGGATTAAAG
 4201 GCAGAAATTT TACATTTCCA ACTTTCCGCT TTTATCTCAA AATCTCAAGG
 4251 TTCACAACCTC GGTAACAGC AAAATGTGGG CGTGAAATTG GGCTATCGTT
 4301 GGTAA

Amino acid sequence for NTHi strain 11 Hap protein (first amino acid to last amino acid):

```

1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENEGKFAVG
51  AKNIDVYNKE GQLVGTSMTK APMIDFSVVS RNGVAALVGD QYIVSVAHNV
101 GYTNVDFGAE GQNPDQHRFT YKIVKRNNYN HDAKHRYLDD YHNPRLHKFV
151 TDAAPIDMTS HMDGNKYANK EKYPERVRVG SGDQYWDDDQ NNRTYLSDGY
201 NYLTGGNTYN QSGRGDGYSY VRGDIRKVG DYGPLPIASSF GDSGSPMFIY
251 DAETQKWLIN GVLREGQPYT GEFDGFQLAR KSFLDEIIRK DQPNGFLTPK
301 GNGVYTISKS DDGIGVVTSK IGKPREIPLA NNKLKIEDKD TVYNNRYNGP
351 NIYSPQLNNG KNIYFGDEEL GSITLTDDID QGAGGLYFEG DFIVSPTKNE
401 TWKGAGIHVS EISTVTWKVN GVENDRLSKI GKGTLHV KAK GENKGSISVG
451 DGKVILEQQA DDQGNKQAFS EIGLVSGRGT VQLNDDKQFD TDKFYFGFRG
501 GRLDLNGHSL TFKRIQNTDE GAMIVNHNTT QVANITITGN ESITAPS NKN
551 NINKLDYSKE IAYNGWFXET DKNKHNGRLN LIYKPTTEDR TLLLSGGTNL
601 KGDITQTKGK LFFSGRPTPH AYNHLDKRWS EMEGIPQGEI VWDYDWINRT
651 FKAENFQIKG GSAVVS RNV SIEGNWTVSN NANATFGVVP NQONTICTRS
701 DWTGLTTCKT VNLTDDKVID SIPTTQINGS INLTDNATVN INGLAKLNGN
751 VTLINHSQFT LSNNATQIGN IKLSNHANAR VNNATLMGDV NLADTSRFTL
801 SNQATQIGTI SLHQQAQATV DNANLNGNVH LTDSARFSLK NSHFSHQIQG
851 DKDTT VTLEN ATWTMPSDTT LQNLTLNNST VTLNSAYSAS SNNAPRRRRS
901 LETETTP TSA EHRFNTLTVN GKLSGQGTFO FTPSLFGYES DKLKLSNDAE
951 GDYTL SVRNT GKEPVTLEQL TLVESKDNKP LSDKLKFTLE NDHVDAGALR
1001 YKLVKNKGEF RLHNPIKEQE LRSDLVRAEQ AERTLEAKQV EQTAETQTSN
1051 ARVRSRR AVL SDTPSAQSLL NAEV KQAE P NAKTQKSKAK TKKARSKRAL
1101 REAFSDTPPD LSQ LNVLEAA LKVINAQPQT EKERQAQEEE AKRQRKQKDL
1151 ISRYSNSALS ELSATVNSML SVQDELDRLF VDQAQSALWT NIAQDKRRYD
1201 SDAFRAYQQK TNL RQIGVQK ALDNGRIGAV FSHSRSDNTF DEQVKNHATL
1251 TMSGFAQYQ WGD LQFGVNV GAGISASKMA EEQSRKIHRK AINYGVNASY
1301 QFRLGQLGIQ PYLGVNRYFI ERENYQSEE V KVQTPSLAFN RYNAGIRVDY
1351 TFTPTDNISV KPYFFVNYVD VSNANVQTTV NSTMLQQSFG RYWQKEVGLK
1401 AEILHFQLSA FISKSQGS QL GKQQNVGVKL GYRW

```

FIG. 17

Nucleotide sequence for NTHi strain TN106 hap gene (start codon begins at position 422, stop codon begins at position 4595):

```

1  TGGCGGCGGA CAAATTATTG CGACGGGTAC ACCAGAACAA GTTGCTAAAG
51  TAAAAAGTTC CCACACCGCT CGCTTCCTTA AACCGATTTT AGAAAAACCT
101 TAGAAAAAAT GACCGCACTT TCAGAGAAAA CTCACATAAA GTGCGGTTAT
151 TTTATTAGTG ATATTGTTTT AATTTTAGTT ATCTGTATAA ATTACATACA
201 ATATTAATCC ATCGCAAGAT TAGATTACCC ACTAAGTATT AAGCAAAAAC
251 CTAGAAATTT TGGCTTAATT ACTATATAGT TTTACTCATT TATTTTCTTT
301 TGTGCCTTTT AGTTCATTTT TTTAGCTGAA ATCCCTTAGA AAATCACCGC
351 ACTTTTATTG TTCAATAGTC GTTTAACCAC GTATTTTTTA ATACGAAAAA
401 TTAATAATT AAATAAACAT TATGAAAAAA ACTGTATTTT GTCTGAATTT
451 TTTAACCGCT TGCATTTTCAT TAGGGATAGT ATCGCAAGCG TGGGCAGGTC
501 ATACTTATTT TGGGATTGAC TACCAATATT ATCGTGATTT TGCCGAGAAT
551 AAAGGGAAGT TTACAGTTGG GGCTCAAGAT ATTGATATCT ACAATAAAAA
601 AGGGGAAATG ATAGGTACGA TGATGAAAGG TGTGCCTATG CCTGATTTAT
651 CTTCCATGGT TCGTGGTGGT TATTCAACAT TGATAAGTGA GCAGCATTTA
701 ATTAGCGTCG CACATAATGT AGGGTATGAT GTCGTTGATT TTGGTATGGA
751 GGGGGAAAAT CCAGACCAAC ATCGTTTTAA GTATAAAGTT GTTAAACGAT
801 ATAATTATAA GAGCGGTGAT AGACAATATA ATGATTATCA ACATCCAAGA
851 TTAGAGAAAT TTGTAACGGA AACTGCACCT ATTGAAATGG TTTTCATATAT
901 GGATGGTAAT CATTACAAA ATTTTAATCA ATATCCTTTG CGAGTTAGAG
951 TTGGAAGTGG GCATCAATGG TGGAAAGACG ATAATAATAA AACCATTGGA
1001 GACTTAGCCT ATGGAGGTTT ATGGTTAATA GGTGGAAATA CCTTTGAAGA
1051 TGGACCAGCT GGTAACGGTA CATTAGAATT AAATGGGCGA GTACAAAATC
1101 CTAATAAATA TGGTCCACTA CCTACGGCAG GTTCATTCCG GGATAGTGGT
1151 TCTCCAATGT TTATTTATGA TAAGGAAGTT AAGAAATGGT TATTAAATGG
1201 CGTGTTACGT GAAGGAAATC CTTATGCTGC AGTAGGAAAC AGCTATCAAA
1251 TTACACGAAA AGATTATTTT CAAGGTATTC TTAATCAAGA CATTACAGCT
1301 AATTTTTGGG ATACTAATGC TGAATATAGA TTTAATATAG GGAGTGACCA
1351 CAATGGAAGA GTGGCAACAA TCAAAAGTAC ATTACCTAAA AAAGCTATTC
1401 AGCCTGAACG AATAGTGGGT CTTTATGATA ATAGCCAAC TCGATGATGCT
1451 AGAGATAAAA ATGGCGATGA ATCTCCCTCT TATAAAGGTC CTAATCCATG
1501 GTCGCCAGCA TTACATCATG GGAAAAGTAT TTAATTTGGC GATCAAGGAA
1551 CAGGAACTTT AACAATTGAA AATAATATAA ATCAAGGTGC AGGTGGATTG
1601 TATTTTGAAG GTAATTTTGT TGTAAGGAGC AATCAAAATA ATATAACTTG
1651 GCAAGGTGCA GCGTTTCTG TTGGAGAAGA AAGTACTGTT GAATGGCAGG
1701 TGCATAATCC AGAAGGCGAT CGCTTATCCA AAATTGGGCT GGGAACCTTA
1751 CTTGTTAATG GTAAAGGGAA AAAGTTAGGA AGCCTGAGTG TCGGTAACGG
1801 TTTGGTTGTG TTAGATCAAC AAGCAGATGA ATCAGGTCAA AAACAAGCCT
1851 TTAAAGAAGT TGGCATTGTA AGTGGTAGAG CTACCGTTCA ACTAAATAGT
1901 GCAGATCAAG TTGATCCTAA CAATATTTAT TTCGGCTTTC GTGGTGGTGC
1951 CTTAGATCTT AATGGGCATT CATTAACCTT TGAACGTATC CAAAATACGG
2001 ATGAAGGCGC GATGATTGTG AACCACAACG CTTCTCAAAC CGCAAATATT

```

FIG. 18A

2051 ACGATTACAG GCAACGCAAC TATTAATTCA GATAGCAAAC AACTTACTAA
2101 TAAAAAAGAT ATTGCATTTA ACGGCTGGTT TGGTGAGCAA GATAAAGCTA
2151 AAACAAATGG TCGTTTAAAT GTGAATTATC AACCAGTTAA TGCAGAAAAT
2201 CATTGTGTC TTTCTGGGGG GACAAATTTA AACGGCAATA TCACGCAAAA
2251 TGGTGGTACG TTAGTTTTTA GTGGTCGTCC AACGCCTCAT GCTTACAATC
2301 ATTTAAGAAG AGACTTGTCT AACATGGAAG GTATCCCACA AGGCGAAATT
2351 GTGTGGGATC ACGATTGGAT CAACCGCACA TTAAAGCTG AAAACTTCCA
2401 AATTAAAGGC GGAAGTGCGG TGGTTTCTCG CAATGTTTCT TCAATTGAGG
2451 GAAATTGGAC AGTCAGCAAT AATGCAAATG CCACATTGG TGTTGTGCCA
2501 AATCAGCAAA ATACCATTGG CACGCGTTCA GATTGGACAG GATTAACGAC
2551 TTGTAAAACA GTTGATTTAA CCGATAAAAA AGTTATTAAT TCCATACCGA
2601 CAACACAAAT TAATGGTTCT ATTAATTTAA CTGATAATGC AACAGTGAAT
2651 ATTCATGGTT TAGCAAACT TAATGGTAAT GTCACTTTAA TAGATCACAG
2701 CCAATTTACA TTGAGCAACA ATGCCACCCA AACAGGCAAT ATCAAACTTT
2751 CAAATCACGC AAATGCAACG GTGGACAATG CAAATTTGAA CGGTAATGTG
2801 AATTTAATGG ATTCTGCTCA ATTTTCTTTA AAAACAGCC ATTTTTCGCA
2851 CCAAATCCAA GGTGGGGAAG ACACAACAGT GATGTTGGAA AATGCGACTT
2901 GGACAATGCC TAGCGATACC ACATTGCAGA ATTTAACGCT AAATAATAGT
2951 ACTGTTACGT TAAATTCAGC TTATTCAGCT ATCTCAAATA ATGCGCCACG
3001 CCGTCGCCGC CGTTCATTAG AGACGGAAC AACGCCAACA TCGGCAGAAC
3051 ATCGTTTCAA CACATTGACA GTAAATGGTA AATTGAGCGG GCAAGGCACA
3101 TTCCAATTTA CTTTCATCTT ATTTGGCTAT AAAAGCGATA AATTAATAAT
3151 ATCCAATGAC GCTGAGGGCG ATTACACATT ATCTGTTTCG AACACAGGCA
3201 AAGAACCCGT GACCTTTGGG CAATTAACCT TGGTTGAAAG CAAAGATAAT
3251 AAACCGTTAT CAGACAACT CACATTCACG TTAGAAAATG ACCACGTTGA
3301 TGCAGGTGCA TTACGTTATA AATTAGTGAA GAATGATGGC GAATTCCGCT
3351 TACATAACCC AATAAAGAG CAGGAATTGC GCTCTGATTT AGTAAGAGCA
3401 GAGCAAGCAG AACGAACATT AGAAGCCAAA CAAGTTGAAC AGACTGCTAA
3451 AACACAAACA AGTAAGGCAA GAGTGCAGTC AAGAAGAGCG GTGTTTTCTG
3501 ATCCCTGCC TGCTCAAAGC CTGTTAAAAG CATTAGAAGC CAAACAAGCT
3551 CTGACTACTG AAACACAAAC AAGTAAGGCA AAAAAAGTGC GGTCAAAAAG
3601 AGCTGCGAGA GAGTTTTCTG ATACCCTGCC TGATCAAATA TTACAAGCCG
3651 CACTTGAGGT TATTGATGCC CAACAGCAAG TGAAAAAGA ACCTCAAACCT
3701 CAAGAGGAAG AAGAGAAAAG ACAACGCAA CAAAAGAAT TGATCAGCCG
3751 TTA CTCAAAT AGTGCGTTAT CGGAGTTGTC TGCGACAGTA AATAGTATGC
3801 TTTCCGTTCA AGATGAATTG GATCGTCTTT TTGTAGATCA AGCACAATCT
3851 GCCGTGTGGA CAAATATCGC ACAGGATAAA AGACGCTATG ATTCTGATGC
3901 GTTCCGTGCT TATCAGCAGA AAACGAACTT GCGTCAAAT GGGGTGCAAA
3951 AAGCCTTAGA TAATGGACGA ATTGGGGCGG TTTTCTCGCA TAGCCGTTCA
4001 GATAATACCT TTGACGAACA GGTAAAAAAT CACGCGACAT TAGCGATGAT
4051 GTCGGGTTT GCCCAATATC AATGGGGCGA TTTACAATTT GGTGTAAACG
4101 TGGGTGCGGG AATTAGTGCG AGTAAAATGG CTGAAGAACA AAGCCGAAAA
4151 ATTCATCGAA AAGCGATAAA TTATGGTGTG AATGCAAGTT ATCAGTTCCG
4201 TTTAGGCAA TTGGGTATTC AGCCTTATTT GGGTGTTAAT CGATATTTTA

FIG.-18B

4251 TTGAACGTGA AAATTATCAA TCTGAAGAAG TGAAAGTGCA AACACCGAGC
 4301 CTTGTATTTA ATCGCTATAA TGCTGGCATT CGAGTTGATT ATACATTTAC
 4351 CCCGACAGAT AATATCAGCA TTAAGCCTTA TTTCTTCGTC AATTATGTTG
 4401 ATGTTTCAAA CGCTAACGTA CAAACCACTG TAAATCGCAC GATGTTGCAA
 4451 CAATCATTTG GCGGTTATTG GCAAAAAGAA GTGGGATTAA AGGCAGAAAT
 4501 TTTACATTTT CAACTTTCCT CTTTTATCTC AAAATCTCAA GGTTCACAAC
 4551 TCGGCAAACA GCAAAATGTG GCGGTGAAAT TGGGGTATCG TTGGTAAAAA
 4601 TCAAC

FIG._18C

Amino acid sequence for NTHi strain TN106 Hap protein (first amino acid to last amino acid):

1 MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
 51 AQDIDIYNKK GEMIGTMMKG VPMPDLSSMV RGGYSTLISE QHLISVAHNV
 101 GYDVVDFGME GENPDQHRFK YKVVKRYNYK SGDRQYNDYQ HPRLEKFBTE
 151 TAPIEMVSYM DGNHYKNFNQ YPLRVRVGSQ HQWWKDDNNK TIGDLAYGGS
 201 WLIGGNTFED GPAGNGTLEL NGRVQNPKNY GPLPTAGSFG DSGSPMFIYD
 251 KEVKKWLLNG VLREGNPYAA VGNSYQITRK DYFQGILNQD ITANFWDNTA
 301 EYRFNIGSDH NGRVATIKST LPKKAIQPER IVGLYDNSQL HDARDKNGDE
 351 SPSYKGPNPW SPALHHGKSI YFGDQGTGTL TIENNINQGA GGLYFEGNFV
 401 VKGNQNNITW QGAGVSVGEE STVEWQVHNP EGDRLSKIGL GTLLVNGKKG
 451 NLGSLSVGNG LVVLDQQADE SGQKQAFKEV GIVSGRATVQ LNSADQVDPN
 501 NIYFGFRGGR LDLNGHSLTF ERIQNTDEGA MIVNHNASQT ANITITGNAT
 551 INSDSKQLTN KKDIAFNGWF GEQDKAKTNG RLNVNYQPVN AENHLLLSGG
 601 TNLNGNITQN GGTLVFSGRP TPHAYNHLRR DLSNMEGIPQ GEIVWDHDWI
 651 NRTFKAENFQ IKGGSVVSR NVSSIEGNWT VSNANATFG VVPNQQNTIC
 701 TRSDWTGLTT CKTVDLTDKK VINSIPTTQI NGSINLTDNA TVNIHGLAKL
 751 NGNVTLIDHS QFTLSNNATQ TGNIKLSNHA NATVDNANLN GNVNLMDSAQ
 801 FSLKNSHFSH QIQGGEDTTV MLENATWTMP SDTTLQNLTN NNSTVTLNSA
 851 YSAISNNAPR RRRRSLETET TPTSAEHRFN TLTVNGKLSG QGTFQFTSSL
 901 FGYKSDKLKL SNDAEGDYTL SVRNTGKEPV TFGQLTLVES KDNKPLSDKL
 951 TFTLENDHVD AGALRYKLK NDGEFRLHNP IKEQELRSDL VRAEQAERTL
 1001 EAKQVEQTAK TQTSKARVRS RRAVFSDDL P AQSLLKALEA KQALTTETQT
 1051 SKAKKVRSKR AAREFSDTLP DQILQAALV IDAQQQVKKE PQTQEEEEKR
 1101 QRKQKELISR YSNSALSELS ATVNSMLSVQ DELDRLFVDQ AQSASVTNIA
 1151 QDKRRYSDA FRAYQQKTNL RQIGVQKALD NGRIGAVFSH SRSDNTFDEQ
 1201 VKNHATLMM SGFAQYQWGD LQFGVNVGAG ISASKMAEEQ SRKIHRKAIN
 1251 YGVNASYQFR LGQLGIQPYL GVNRYFIERE NYQSEEVKVQ TPSLVFNRYN
 1301 AGIRVDYFTT PTDNISIKPY FFFVNYVDVSN ANVQTTVNRT MLQQSFGRYW
 1351 QKEVGLKAEI LHFQLSAFIS KSQGSQLGKQ QNVGVKLGYS W

FIG._19

Nucleotide sequence for NTHi strain 860295 *hap* gene (start codon begins at position 430, stop codon begins at position 4738):

```

1  GGAGGCAGTG GTGGCGGACA AATTATTGCG ACGGGTACGC CAGAACAAGT
51  TGCCAAAGTA GAAAGTTCCC ACACCGCCCG CTTCTTAAA CCGATTTTAG
101 AAAAACCTTA GAAAAAATGA CCGCACTTTC AGAGAAAACT CACATAAAGT
151 GCGGTTATTT TATTAGTGAT ATTGTTTTAA TTTTAGTTAT CTGTATAAAT
201 TACATATAAT ATTAATCCAT CGCAAGATAA GATTACCCAC TAAGTATTAA
251 GCAAAAACCT AGAAATTTTG GCTTAATTAC TATATAGTTT TACTGCTTTA
301 TTTTCTTTTG TGCCTTTTAG TTCGTTTTTT TAGCTGAAAT CCCTTAGAAA
351 ATCACCGCAC TTTTATTGTT CAATAGTCGT TTAACCACGT ATTTTTTAAT
401 ACGAAAAATT ACTTAATTAA ATAAACATTA TGAAAAAAC TGTATTTTCG
451 CTGAACTTTT TAACCGCTTG CATTTCATTG GGGATAGTAT CGCAAGCGTG
501 GGCAGGTCAC ACTTATTTTG GGATTGACTA CCAATATTAT CGTGATTTTG
551 CTGAGAATAA AGGGAAGTTT TCAGTTGGGG CTAAAAATAT TGAGGTTTAT
601 AACAAAGAGG GGACTTTAGT TGGCACATCA ATGACAAAAG CCCCATGAT
651 TGATTTTTCT GTGGTGTGCG GAAATGGGGT GGCGGCATTA GTAGGCGATC
701 AGTATATTGT GAGTGTGGCA CATAACGGTG GATATAATAG CGTTGATTTT
751 GGAGCAGAAG GTCCAAATCC CGATCAGCAT CGTTTTACTT ATCAAATTGT
801 AAAAAGAAAT AATTATAAGC CAGGCAAAGA TAACCCTTAT CATGGTGAAT
851 ATCACATGCC TCGTTTGCAC AAATTTGTCA CTGACGCTGA ACCAGCAAAG
901 ATGACAGACA ATATGAATGG AAAGAACTAC GCTGATTTAA GTAAATATCC
951 TGATCGTGTG CGTATTGGTA CAGGTGAACA ATGGTGGAGG ACTGATGAAG
1001 AACAAAAGCA AGGAAGTAAG AGTTCATGGC TTGCTGATGC TTATCTGTGG
1051 AGAATAGCAG GTAACACACA TTCACAAAGT GGAGCGGGCA ACGGCACGGT
1101 AAACCTAAGT GGAGATATCA CAAAACCAA TAACATATGA CCTCTTCTTA
1151 CGGGTGTTTC GTTTGGAGAT AGTGGTTCTC CAATGTTTAT TTATGATGCA
1201 ATAAAACAAA AATGGCTTAT TAATGGCGTA TTGCAAACTG GTAACCTTTT
1251 CTCGGGAGCT GGAAATGGAT TCCAATTAAT TAGAAAAAAT TGGTTTTATG
1301 ATAATGTCTT TGTAGAAGAT TTGCCTATAA CATTTTTAGA GCCAAGAAGT
1351 AACGGTCATT ATTCATTTAC TTCAAATAAT AATGGAAGT GTACGGTTAC
1401 TCAAACGAAT GAAAAAGTGA GTATGCCTCA ATTTAAAGTC AGAACGGTTC
1451 AGTTATTTAA TGAAGCATTG AAAGAAAAAG ATAAAGAACC TGTTTTATGCT
1501 GCAGGTGGTG TAAATGCTTA TAAACCAAGA CTAAATAATG GTAAAAATAT
1551 TTACTTTGGC GATCGAGGAA CAGGAAGTTT AACAATTGAA AATAATATAA
1601 ATCAAGGTGC TGGTGGTTTG TATTTTGAGG GTAACCTTAC GGTATCTTCA
1651 GAAAATAATG CAACTTGGCA AGGTGCTGGA GTGCATGTAG GTGAAGACAG
1701 TACTGTTACT TGGAAAGTAA ACGGCGTGGA ACATGATCGC CTTTCTAAAA
1751 TTGGTAAAGG AACGTTGCAT ATTCAAGCAA AAGGTGAAAA CTTAGGCTCA
1801 ATTAGCGTAG GTGACGGCAA AGTCATTTTA GATCAACAAG CCGATGAGAA
1851 CAACCAAAAA CAAGCCTTTA AAGAAGTTGG CATTGTAAGT GGTAGAGCTA
1901 CCGTTCAACT AAATAGTGCA GATCAAGTTG ATCCTAACAA TATTTATTTT
1951 GGATTTCTGT GTGGTCGCTT AGATCTTAAC GGACATTCAT TAACCTTTAA
2001 ACGTATCCAA AATACGGACG AGGGCGCGAT GATTGTGAAC CATAATACAA

```

FIG._20A

2051 CTCAAGTCGC TAATATTACT ATTACTGGGA ACGAAAGTAT TACTGCTCCA
2101 TCTAATAAAA ATAATATTAA TAAACTTGAT TACAGCAAAG AAATTGCTTA
2151 CAACGGTTGG TTTGGCGAAA CAGATGAAAA TAAACACAAT GGAAGATTAA
2201 ACCTTATTTA TAAACCAACC ACAGAAGATC GTACTTTGCT ACTTTCAGGT
2251 GGAACAAATT TAAAAGGCAA TATTACTCAG GAAGGCGGCA CTTTAGTGTT
2301 TAGTGGTCGC CCAACTCCAC ACGCTTACAA TCATTTAAAT CGCCCAAACG
2351 AGCTTGGGCG ACCTCAAGGC GAAGTGGTTA TTGATGACGA TTGGATCACC
2401 CGCACATTTA AAGCTGAAAA CTTCCAAATT AAAGGCGGAA GTGCGGTGGT
2451 TTCTCGCAAT GTTCTTTCAA TTGAGGGAAA TTGGACAGTC AGCAATAATG
2501 CAAATGCCGC ATTTGGTGTT GTGCCAAATC AGCAAATAAC CATTTGCACG
2551 CGTTCAGATT GGACAGGATT AACGACTTGT AAAACTGTGG ATTTAACCGA
2601 TACAAAAGTT ATTAATTCCA TACCGACAAC ACAAATTAAT GGCTCTATTA
2651 ATTTAACTGA TAATGCAACA GTGAATATTC ATGGTTTAGC AAAACTTAAT
2701 GGTAATGTCA CTTTAATAAA TCATAGCCAA TTTACATTGA GCAACAATGC
2751 CACCCAAACA GGCAATATCC AACTTTCAAA TCACGCAAAT GCAACGGTGG
2801 ACAATGCAAA TTTGAACGGT AATGTGCATT TAACGGATTG TGCTCAATTT
2851 TCTTTAAAAA ACAGCCATTT TTCGCACCAA ATTCAGGGCG ACAAAGACAC
2901 AACAGTGACG TTGGAAAATG CGACTTGGAC AATGCCTAGC GATGCCACAT
2951 TGCAGAATTT AACGCTAAAT AATAGTACTG TTACGTAAAT TTCAGCTTAT
3001 TCAGCTAGCT CAAATAATGC GCCACGTCAC CGCCGTTCAT TAGAGACGGA
3051 AACACGCCA ACATCGGCAG AACATCGTTT CAACACATTG ACAGTAAATG
3101 GTAAATTGAG CGGGCAAGGC ACATTCCAAT TTA CTTCATC TTTATTTGGC
3151 TATAAAAGCG ATAAATTAAA ATTATCCAAT GACGCTGAGG GCGATTACAC
3201 ATTATCTGTT CGCAACACAG GCAAAGAACC CGAAGCCCTT GAGCAATTA
3251 CTTTGGTTGA AAGCAAAGAT AATAAACCGT TATCAGACAA ACTCAAATTT
3301 ACTTTAGAAA ATGACCACGT TGATGCAGGT GCATTACGTT ATAAATTAGT
3351 GAAGAATAAT GGCGAATTCC GCTTGATAA CCCAATAAAA GAGCAGGAAT
3401 TGCGCAATGA TTTAGTAAGA GCAGAGCAAG CAGAACGAAC ATTAGAAGCC
3451 AAACAAGTTG AACAGACTGC TGAAACACAA ACAAGTAATG CAAGAGTGCG
3501 GTCAAAAAGA GCGGTGTTTT CTGATACCCT GCCTGATCAA AGCCAGTTAG
3551 ACGTATTACA AGCCGAACAA GTTGAACCGA CTGCTGAAAA ACAAAAAAAT
3601 AAGGCAAAAA AAGTGCGGTC AAAAAGAGCG GTGTTTTCTG ATACCCTGCC
3651 TGATCAAAGC CAGTTAGACG TATTACAAGC CGAACAAGTT GAACCGACTG
3701 CTGAAAAACA AAAAAATAAG GCAAAAAAAG TGCCGTCAAA AAGAGCCGCG
3751 AGAGAGTTTT CTGATACCCC GCTTGATCTA AGCCGGTTAA AGGTATTAGA
3801 AGTCAAACCT GAGGTATTAT ATGCCCAACA GCAAGTGAAA AAAGAACCTC
3851 AAGATCAAGA GAAACAACGC AAACAAAAAG ACTTGATCAG CCGTTATTCA
3901 AATAGTGCGT TATCAGAATT ATCTGCAACA GTAAATAGTA TGCTTTCTGT
3951 TCAAGATGAA TTAGATCGTC TTTTGTAGA TCAAGCACAA TCTGCCGTGT
4001 GGACAAATAT CGCACAGGAT AAAAGACGCT ATGATTCTGA TCGGTTCCGT
4051 GCTTATCAGC AGAAAACGAA CTTACGTCAA ATTGGGGTGC AAAAAGCCTT
4101 AGCTAATGGA CGAATTGGGG CAGTTTTCTC GCATAGCCGT TCAGATAATA
4151 CTTTTGATGA ACAGGTTAAA AATCACGCGA CATTAAACGAT GATGTCGGGT
4201 TTTGCCCAAT ATCAATGGGG CGATTTACAA TTTGGTGTA ACGTGGGAAC

FIG._20B

4251 GGGAAATCAGT GCGAGTAAAA TGGCTGAAGA ACAAAGCCGA AAAATTCATC
4301 GAAAAGCGAT AAATTATGGC GTGAATGCAA GTTATCAGTT CCGTTTAGGG
4351 CAATTGGGCA TTCAGCCTTA TTTTGGAGTT AATCGCTATT TTATTGAACG
4401 TGAAAATTAT CAATCTGAGG AAGTGAAAGT GAAAACGCCT AGCCTTG CAT
4451 TTAATCGCTA TAATGCTGGC ATTCGAGTTG ATTATACATT TACTCCGACA
4501 GATAATATCA GCGTTAAGCC TTATTTCTTC GTCAATTATG TTGATGTTTC
4551 AAACGCTAAC GTACAAACCA CGGTAAATAG CACGGTGTTG CAACAACCAT
4601 TTGGACGTTA TTGGCAAAAA GAAGTGGGAT TAAAAGCGGA AATTTTACAT
4651 TTCCAACTTT CTGCTTTTAT TTCTAAATCT CAAGGTTCGC AACTCGGCAA
4701 ACAGCAAAAT GTGGGCGTGA AATTGGGGTA TCGTTGGTAA AAATCAACAT
4751 AATTGTATCG TTTATTGATA AACAAGGTGG GGCAGATCCC ACCTTTTTTA
4801 TTTCAATAAT GGAACTTTAT TTAATTAAGA GCATCTAAGT AGCACCCCAT
4851 ATAGGGGATT AATTAAGAGG ATTTAATAAT GAATTTAACT AAACTTTAC
4901 CAGCATTTGC TGCTGCAGTC GTATTATCTG CTTGTGCAAA GGATGCACCT
4951 GAAATGACAA AATCATCTGC GCAAATAGCT GAAATGCAAA CACTTCCAAC
5001 AATCACTGAT AAAACAGTTG TATATTCCTG CAATAAACAA ACTGTAACTG
5051 CCGTGTATCA ATTTGAAAAC CAAGAACCAG TTGCTGCAAT GGTAAGTGTG
5101 GGCGATGGCA TTATTGCGAA AGATTTTACT CGTGATAAAT CACAAAATGA
5151 CTTTACAAGT TTCGTTTCTG GGGATTATGT TTGGAATGTA GATAGTGGCT
5201 TAACGTTAGA TAAATTTGAT TCTGTTGTGC CTGTCAATTT AATTC

FIG._20C

Amino acid sequence for NTHi strain 860295 Hap protein (first amino acid to last amino acid):

```

1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFSVG
51  AKNIEVYNKE GTLVGTSMTK APMIDFSVVS RGVAAALVGD QYIVSVAHNG
101 GYNSVDFGAE GPNPDQHRFT YQIVKRNNYK PGKDNPHYGD YHMPRLHKFV
151 TDAEPAKMTD NMNGKNYADL SKYPDRVRIG TGEQWVRTDE EQKQGSKSSW
201 LADAYLWRIA GNTHSQSGAG NGTVNLSGDI TKPNNYGPLP TGVSFSGDSGS
251 PMFIYDAIKQ KWLINGVLQT GNPFSGAGNG FQLIRKNWFY DNVFVEDLPI
301 TFLEPRSNHG YSFTSNNNGT GTVTQTNEKV SMPQFKVRTV QLFNEALKEK
351 DKEPVYAAGG VNAYKPRLNN GKNIYFGDRG TGTLTIENTNI NQGAGGLYFE
401 GNFTVSSENN ATWQGAGVHV GEDSTVTWKV NGVEHDLRSK IGKGTLHIQA
451 KGENLGSISV GDGKVILDQQ ADENNQKQAF KEVGIVSGRA TVQLNSADQV
501 DPNNIYFGFR GGRDLNLGHS LTFKRIQNTD EGAMIVNHNT TQVANITITG
551 NESITAPSNK NNINKLDYSK EIAYNGWFGE TDENKHNGRL NLIYKPTTED
601 RTLLLSGGTN LKGNITQEGG TLVFSGRPTP HAYNHLNRPN ELGRPQGEVV
651 IDDDWITRTF KAENFQIKGG SAVVSRNVSS IEGNWTVSNN ANAAGVVPN
701 QQNTICTRSD WTGLTTCKTV DLTDTKVINS IPTTQINGSI NLTDNATVNI
751 HGLAKLNGNV TLINHSQFTL SNNATQTGNI QLSNHANATV DNANLNGNVH
801 LTDSAQFSLK NSHFSHQIQG DKDTTVTLEN ATWTMPSDAT LQNLTLNNST
851 VTLNSAYSAS SNNAPRHRRS LETETTP TSA EHRFNTLTVN GKLSGQGTFO
901 FTSSLFGYKS DKLKLSNDAE GDYTLNVRNT GKEPEALEQL TLVESKDNKP
951 LSDKLKFTLE NDHVDAGALR YKLVKNNGEF RLHNPIKEQE LRNDLVRAEQ
1001 AERTLEAKQV EQTAETQTSN ARVRSKRAVF SDTLPDQSQL DVLQAEQVEP
1051 TAEKQKNKAK KVRSKRAVFS DTLPDQSQLD VLQAEQVEPT AEKQKNKAKK
1101 VRSKRAAREF SDTPLDL SRL KVLEVKLEVI NAQQQVKKEP QDQEKQRKQK
1151 DLISRYNSA LSELSATVNS MLSVQDELDR LFVDQAQSAV WTNIAQDKRR
1201 YSDAFRAYQ QKTNLRLQIGV QKALANGRIG AVFHSRSDN TFDEQVKNHA
1251 TLTMMSGFAQ YQWGD LQFGV NVGTGISASK MAEEQSRKIH RKAINYGVNA
1301 SYQFRLGQLG IQPYFGVNRY FIERENYQSE EVKVKTPLA FNRYNAGIRV
1351 DYTFTPTDNI SVKPYFFVNY VDVSANANVQT TVNSTVLQQP FGRYWQKEVG
1401 LKAEILHFQL SAFISKSGS QLGKQQNVGV KLG YRW

```

FIG._21

Nucleotide sequence for NTHi strain 3219B *hap* gene (start codon begins at position 388, stop codon begins at position 4561):

```

1  CCTGAAGACG TTGCTCAAGT TAAAGGCTCT CACACAGCCC GATTCCTTAA
51  ACCGATTTTA GAAAAACCTT AGAAAAAATG ACCGCACTTT CAGAGAAAAC
101 TCACATAAAG TGCGGTATT TTATTAGTGA TATTGTTTTA ATTATTTGTA
151 TAAATTACAT ACAATATTAA TCCATCGAAA AATAAGATTA CCCACTAAGT
201 ATTAAGCCAA AACCTAGAAA TTTTGGCTTA ATTACTATAT AATTTTACTC
251 CTTTATTTTC TTTTGTGCCT TTTAGTTAGT TCGTTTTTTA GCTGAAATCC
301 CTCAGAAAAT CACCGCACTT TTATTGTTCA ATAGTCGTTT AACCACGTAT
351 TTTTAAATAC GAAAAATTAC TTAATTAAAT AAACATTATG AAAAAAATG
401 TATTTTCGTCT TAATTTTCTA ACCGCTTGTA TTTCATTAGG GATAGTATCG
451 CAAGCGTGGG CAGGTCACAC TTATTTTGGG ATTGACTACC AATATTATCG
501 TGATTTTGCC GAGAATAAAG GGAAGTTTAC AGTTGGGGCT CAAGATATTG
551 ATATCTACAA TAAAAAAGGG GAAATGATAG GTACGATGAT GAAAGGTGTG
601 CCTATGCCTG ATTTATCTTC CATGGTTCGT GGTGGTTATT CAACATTGAT
651 AAGTGAGCAG CATTTAATTA GCGTCGCACA TAATGTAGGG TATGATGTCG
701 TTGATTTTGG TATGGAGGGG GAAAATCCAG ACCAACATCG TTTTAAGTAT
751 AAAGTTGTTA AACGATATAA TTATAAGAGC GGTGATAGAC AATATAATGA
801 TTATCAACAT CCAAGATTAG AGAAATTGT AACGGAACT GCACCTATTG
851 AAATGGTTTC ATATATGGAT GGTAATCATT ACAAAAATTT TAATCAATAT
901 CCTTTGCGAG TTAGAGTTGG AAGTGGGCAT CAATGGTGGA AAGACGATAA
951 TAATAAAACC ATTGGAGACT TAGCCTATGG AGGTTCATGG TTAATAGGTG
1001 GAAATACCTT TGAAGATGGA CCAGCTGGTA ACGGTACATT AGAATTAAAT
1051 GGGCGAGTAC AAAATCCTAA TAAATATGGT CCACTACCTA CGGCAGGTTC
1101 ATTCGGGGAT AGTGGTTCTC CAATGTTTAT TTATGATAAG GAAGTTAAGA
1151 AATGGTTATT AAATGGCGTG TTACGTGAAG GAAATCCTTA TGCTGCAGTA
1201 GGAAACAGCT ATCAAATTAC ACGAAAAGAT TATTTTCAAG GTATTCTTAA
1251 TCAAGACATT ACAGCTAATT TTTGGGATAC TAATGCTGAA TATAGATTTA
1301 ATATAGGGAG TGACCACAAT GGAAGAGTGG CAACAATCAA AAGTACATTA
1351 CCTAAAAAAG CTATTCAGCC TGAACGAATA GTGGGTCTTT ATGATAATAG
1401 CCAACTTCAT GATGCTAGAG ATAAAAATGG CGATGAATCT CCCTCTTATA
1451 AAGGTCCTAA TCCATGGTCG CCAGCATTAC ATCATGGGAA AAGTATTTAC
1501 TTTGGCGATC AAGGAACAGG AACTTTAACA ATTGAAAATA ATATAAATCA
1551 AGGTGCAGGT GGATTGTATT TTGAAGGTAA TTTTGTGTGA AAAGGCAATC
1601 AAAATAATAT AACTTGGCAA GGTGCAGGCG TTTCTGTTGG AGAAGAAAGT
1651 ACTGTTGAAT GGCAGGTGCA TAATCCAGAA GGCGATCGCT TATCCAAAAT
1701 TGGGCTGGGA ACCTTACTTG TTAATGGTAA AGGGAAAAAC TTAGGAAGCC
1751 TGAGTGTCGG TAACGGTTTG GTTGTTAGT ATCAACAAGC AGATGAATCA
1801 GGTCAAAAAC AAGCCTTTAA AGAAGTTGGC ATTGTAAGTG GTAGAGCTAC
1851 CGTTCAACTA AATAGTGCAG ATCAAGTTGA TCCTAACAAT ATTTATTTTCG
1901 GCTTTCGTGG TGGTCGCTTA GATCTTAATG GGCATTCAAT AACCTTTGAA
1951 CGTATCCAAA ATACGGATGA AGGCGCGATG ATTGTGAACC ACAACGCTTC
2001 TCAAACCGCA AATATTACGA TTACAGGCAA CGCAACTATT AATTCAGATA

```


2051 GCAAACAAC TACTAATAAA AAAGATATTG CATTTAACGG CTGGTTTGGT
 2101 GAGCAAGATA AAGCTAAAAC AAATGGTCGT TTAAATGTGA ATTATCAACC
 2151 AGTTAATGCA GAAAATCATT TGTTGCTTTC TGGGGGGACA AATTTAAACG
 2201 GCAATATCAC GCAAAATGGT GGTACGTTAG TTTT TAGTGG TCGTCCAACG
 2251 CCTCATGCTT ACAATCATTT AAGAAGAGAC TTGTCTAACA TGGAAGGTAT
 2301 CCCACAAGGC GAAATTGTGT GGGATCACGA TTGGATCAAC CGCACATTTA
 2351 AAGCTGAAAA CTTCCAAATT AAAGGCGGAA GTGCGGTGGT TTCTCGCAAT
 2401 GTTTCTTCAA TTGAGGGAAA TTGGACAGTC AGCAATAATG CAAATGCCAC
 2451 ATTTGGTGTT GTGCCAAATC AGCAAATAAC CATTTGCACG CGTTCAGATT
 2501 GGACAGGATT AACGACTTGT AAAACAGTTG ATTTAACCGA TAAAAAAGTT
 2551 ATTAATTCCA TACCGACAAC ACAAATTAAT GGTTCTATTA ATTTAACTGA
 2601 TAATGCAACA GTGAATATTC ATGGTTTAGC AAAACTTAAT GGTAATGTCA
 2651 CTTTAATAGA TCACAGCCAA TTTACATTGA GCAACAATGC CACCCAAGCA
 2701 GGCAATATCA AACTTTCAAA TCACGCAAAT GCAACGGTGG ACAATGCAAA
 2751 TTTGAACGGT AATGTGAATT TAATGGATTG TGCTCAATTT TCTTTAAAAA
 2801 ACAGCCATTT TTCGCACCAA ATCCAAGGTG GGGAAGACAC AACAGTGATG
 2851 TTGGAAAATG CGACTTGGAC AATGCCTAGC GATACCACAT TGCAGAATTT
 2901 AACGCTAAAT AATAGTACTG TTACGTTAAA TTCAGCTTAT TCAGCTATCT
 2951 CAAATAATGC GCCACGCCGT CGCCGCCGTT CATTAGAGAC GGAAACAACG
 3001 CCAACATCGG CAGAACATCG TTTCAACACA TTGACAGTAA ATGGTAAATT
 3051 GAGCGGGCAA GGCACATTCC AATTTACTTC ATCTTTATTT GGCTATAAAA
 3101 GCGATAAATT AAAATTATCC AATGACGCTG AGGGCGATTA CACATTATCT
 3151 GTTCGCAACA CAGGCAAAGA ACCCGTGACC TTTGGGCAAT TAACTTTGGT
 3201 TGAAAGCAAA GATAATAAAC CGTTATCAGA CAAACTCACA TTCACGTTAG
 3251 AAAATGACCA CGTTGATGCA GGTGCATTAC GTTATAAATT AGTGAAGAAT
 3301 GATGGCGAAT TCCGCTTACA TAACCCAATA AAAGAGCAGG AATTGCGCTC
 3351 TGATTTAGTA AGAGCAGAGC AAGCAGAACG AACATTAGAA GCCAAACAAG
 3401 TTGAACAGAC TGCTAAAACA CAAACAAGTA AGGCAAGAGT GCGGTCAAGA
 3451 AGAGCGGTGT TTTCTGATCC CCTGCCTGCT CAAAGCCTGT TAAACGCATT
 3501 AGAAGCCAAA CAAGCTCTGA CTACTGAAAC ACAAACAAGT AAGGCAAAAA
 3551 AAGTGCGGTC AAAAAGAGCT GCGAGAGAGT TTTCTGATAC CCTGCCTGAT
 3601 CAAATATTAC AAGCCGCACT TGAGGTTATT GATGCCCAAC AGCAAGTGAA
 3651 AAAAGAACCT CAAACTCAAG AGGAAGAAGA GAAAAGACAA CGCAAACAAA
 3701 AAGAATTGAT CAGCCGTTAC TCAAATAGTG CGTTATCGGA GTTGTCTGCG
 3751 ACAGTAAATA GTATGCTTTC CGTTCAAGAT GAATTGGATC GTCTTTTTGT
 3801 AGATCAAGCA CAATCTGCCG TGTGGACAAA TATCGCACAG GATAAAAGAC
 3851 GCTATGATTC TGATGCGTTC CGTGCTTATC AGCAGAAAAC GAACTTGCGT
 3901 CAAATTGGGG TGCAAAAAGC CTTAGATAAT GGACGAATTG GGGCGGTTTT
 3951 CTCGCATAGC CGTTCAGATA ATACCTTTGA CGAACAGGTT AAAAATCACG
 4001 CGACATTAGC GATGATGTCT GGTTTTGCCC AATATCAATG GGGCGATTTA
 4051 CAATTTGGTG TAAACGTGGG TGCGGGAATT AGTGCGAGTA AAATGGCTGA
 4101 AGAACAAAGC CGAAAAATTC ATCGAAAAGC GATAAATTAT GGTGTGAATG
 4151 CAAGTTATCA GTTCCGTTTA GGGCAATTGG GTATTCAGCC TTATTGGGT
 4201 GTTAATCGAT ATTTTATTGA ACGTGAAAAT TATCAATCTG AAGAAGTGAA

```

4251 AGTGCAAACA CCGAGCCTTG TATTTAATCG CTATAATGCT GGCATTTCGAG
4301 TTGATTATAC ATTTACCCCG ACAGATAATA TCAGCATTAA GCCTTATTTTC
4351 TTCGTCAATT ATGTTGATGT TTCAAACGCT AACGTACAAA CCACTGTAAA
4401 TCGCACGATG TTGCAACAAT CATTGCGGCG TTATTGGCAA AAAGAAGTGG
4451 GATTAAAGGC AGAAATTTTA CATTTCACAC TTTCCGCTTT TATCTCAAAA
4501 TCTCAAGGTT CACAACCTCG CAAACAGCAA AATGTGGGCG TGAAATTGGG
4551 GTATCGTTGG TAAAAATCAA CATAATTTTA TCGTTTATTG ATAAACAAGG
4601 TGGGGCAGAT CAAATCCTAC CTTTTTTTATT CCAATAATGG AACTTTATTT
4651 TATTAAAGGT ATCTAAGTAG CACCCTATAT AGGGATTAAT TAAGAGGATT
4701 TAATAATGAA TTAACTAAA ATTTTACCCA CATTGCTGCTG TGTAGTCGTA
4751 TTATCTGCTT GTGCAAAGGA TGCACCTGAA ATGACAAAAT CATCTGCGCA
4801 AATAGCTGAA ATGCAAACAC TT

```

FIG._22C

Amino acid sequence for NTHi strain 3219B Hap protein (first amino acid to last amino acid):

```

1  MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
51  AQDIDIYNKK GEMIGTMMKG VPMPDLSSMV RGGYSTLISE QHLISVAHNV
101 GYDVVDFGME GENPDQHRFK YKVVKRYNYK SGDRQYNDYQ HPRLEKFTVE
151 TAPIEMVSYM DGNHYKNFNQ YPLRVRVGSG HQWWKDDNNK TIGDLAYGGS
201 WLIGGNTFED GPAGNGTLEL NGRVQNPKNKY GPLPTAGSFG DSGSPMFIYD
251 KEVKKWLLNG VLREGNYPYAA VGNSYQITRK DYFQGILNQD ITANFWDNTA
301 EYRFNIGSDH NGRVATIKST LPKKAIQPER IVGLYDNSQL HDARDKNGDE
351 SPSYKGPNPW SPALHHGKSI YFGDQGTGTL TIENNINQGA GGLYFEGNFV
401 VKGNQNNITW QGAGVSVGEE STVEWQVHNP EGDRLSKIGL GTLLVNGKKG
451 NLGSLSVGNG LVVLDQQADE SGQKQAFKEV GIVSGRATVQ LNSADQVDPN
501 NIYFGFRGGR LDLNGHSLTF ERIQNTDEGA MIVNHNASQT ANITITGNAT
551 INSDSKQLTN KKDIAFNGWF GEQDKAKTNG RLNVNYQPVN AENHLLLSGG
601 TNLNGNITQN GGTLVFSGRP TPHAYNHLRR DLSNMEGIPQ GEIVWDHDWI
651 NRTFKAENFQ IKGGSVVSR NVSSIEGNWT VSNANATFG VVPNQQNTIC
701 TRSDWTGLTT CKTVDLTDKK VINSIPTQI NGSINLTDNA TVNIHGLAKL
751 NGNVTLIDHS QFTLSNNATQ AGNIKLSNHA NATVDNANLN GNVNLMDSAQ
801 FSLKNSHFSH QIQGGEDTTV MLENATWTMP SDTTLQNLTL NNSTVTLNSA
851 YSAISNNAPR RRRRSLETET TPTSAEHRFN TLTVNGKLSG QGTFQFTSSL
901 FGYKSDKLKL SNDAEGDYTL SVRNTGKEPV TFGQLTLVES KDNKPLSDKL
951 TFTLENDHVD AGALRYKLK V NDGEFRLHNP IKEQELRSDL VRAEQAERTL
1001 EAKQVEQTAK TQTSKARVRS RRAVFSDDL AQSLLNALEA KQALTTETQT
1051 SKAKKVRSKR AAREFSDTLP DQILQAALEV IDAQQQVKKE PQTQEEEEKR
1101 QRKQKELISR YSNSALSELS ATVNSMLSVQ DELDRLFVDQ AQSASVTNIA
1151 QDKRRYSDA FRAYQQKTNL RQIGVQKALD NGRIGAVFSH SRSDNTFDEQ
1201 VKNHATLAMB SGFAQYQWGD LQFGVNVGAG ISASKMAEEQ SRKIHRKAIN
1251 YGVNASYQFR LGQLGIQPYL GVNRYFIERE NYQSEEVKVQ TPLSVFNRYN
1301 AGIRVDYFTT PTDNISIKPY FVNVYVDVSN ANVQTTVNRT MLQQSFGRYW
1351 QKEVGLKAEI LHFQLSAFIS KSQGSQGLGKQ QNVGVKLGYS W

```

FIG._23

Nucleotide sequence for NTHi strain 1396B hap gene (start codon begins at position 313, stop codon begins at position 4546):

```

1  TGACCGCACT TTCAGAGAAA ACTCACATAA AGTGCGGTTA TTTTATTAGT
51  GATATTGTTT TAATTTTAGT TATCTGTATA AATTACATAC AATATTAATC
101 CATCGCAAGA TAAGATTACC CACTAAGTAT TAAGCAAAAA CCTAGAAATT
151 TTGGCTTAAT TACTATATAG TTTTACTCAT TTATTTTCTT TTGTGCCTTT
201 TAGTTCGTTT TTTTAGCTGA AATCCCTTAG AAAATCACCG CACTTTTATT
251 GTTCAATAGT CGTTTAACCA CGTATTTTTT AATACGAAAA ATTACTTAAT
301 TAAATAAACA TTATGAAAAA AACTGTATTT CGTCTGAATT TTTTAACCGC
351 TTGCATTTCA TTAGGGATAG TATCGCAAGC GTGGGCAGGT CATACTTATT
401 TTGGGATTGA CTACCAATAT TATCGTGATT TTGCCGAGAA TAAAGGGAAG
451 TTCACAGTTG GGGCTAAAAA TATTGAGGTT TACAATAAAA ATGGAAATTT
501 AGTTGGCACA TCAATGACAA AAGCCCCAAT GATTGATTTT TCCGTGGTGT
551 CGCGAAATGG GGTGGCGGCA TTGGTGGGCG ATCAGTATAT TGTGAGTGTG
601 GCACATAATG TAGGCTATAC CAATGTGGAT TTTGGTGCTG AAGGACAAAA
651 TCCTGATCAA CATCGTTTTA CTTATAAAAT TGTGAAACGG AATAATTATA
701 AAAACGATCA AACGCATCCT TATGAGAAAG ACTACCACAA CCCACGCTTA
751 CATAAATTTG TTACGGAAGC CACCCCAATC GATATGACTT CTGATATGAA
801 CGGCAACAAA TATACAGATA GGACGAAATA TCCCGAACGC GTGCGTATCG
851 GCTCCGGGTG GCAGTTTTTG CGAAACGATC AAAACAACGG CGACCAAGTT
901 GCCGGCGCAT ATCATTACCT GACAGCAGGC AATACACACA ACCAAGGCGG
951 AGCAGGGGGC GGCTGGTCAA GTCTGAGCGG CGATGTGCGC CAAGCGGGCA
1001 ATTACGGCCC CATTCCTATT GCAGGCTCAA GCGGCGACAG CGGTCGCCT
1051 ATGTTTATTT ATGATGCGGA AAAACAAAAA TGGTTGATTA ACGGCGTATT
1101 GAGGACCGGC AACCCTTGGG CGGGGACAGA GAATACATTC CAACTGGTAC
1151 GCAAGTCTTT TTTTGATGAA ATCCTTGAAA AAGATTTGCG TACATCGTTT
1201 TATAGCCCAT CGGGCAATGG TGCATACACC ATTACAGACA AAGGCGACGG
1251 CAGCGGCATT GTCAAACAAC AAACAGGAAG ACCATCTGAA GTCCGCATCG
1301 GTTTAAAAGA CGACAAATTA CCTGCCGAAG GTAAAGACGA TGTTTACCAA
1351 TACCAAGGTC CAAATATATA CCTGCCTCGT TTGAATAACG GTGGAAACCT
1401 GTATTTCCGA GATCAAAAAA ACGGCACTGT TACCTTATCA ACCAACATCA
1451 ACCAAGGTGC GGGCGGTTTG TATTTTGAGG GTAACTTTAC GGTATCTTCA
1501 GAAAATAATG CAACTTGGCA AGGTGCTGGA GTGCATGTAG GTGAAGACAG
1551 TACTGTTACT TGGAAAGTAA ATGGTGTTGA AAATGATCGC CTTTCTAAAA
1601 TCGGCAAAGG CACATTGCAC GTTAAAGCCA AAGGGGAAAA TAAAGGTTTG
1651 ATCAGCGTAG GCGATGGTAA AGTCATTTTG GAGCAGCAGG CAGACGATCA
1701 AGGCAACAAA CAAGCCTTTA GTGAAATTGG CTTGGTTAGT GGCAGAGGTA
1751 CGGTTTCAGT AAACGATGAC AAGCAATTTA ATACTGATAA ATTTTATTTT
1801 GGCTTCCGTG GTGGTCGCTT AGATCTTAAT GGGCATT CAT TAACCTTTAA
1851 ACGTATCCAA AATACGGATG AGGGAGCAAC GATTGTTAAT CACAATGCCA
1901 CAACAGAATC TACAGTGACC ATTACTGGCA GCGATACCAT TAATGACAAC
1951 ACTGGCGATT TAACCAATAA ACGTGATATT GCTTTTAATG GTTGGTTTGG
2001 TGATAAAGAT GATACTAAAA ATACTGGACG TTTGAATGTT ACTTACAATC

```

2051 CGCTTAACAA AGATAATCAC TTCCTTCTAT CAGGTGGAAC AAATTTAAAA
 2101 GGCAATATTA CTCAAGACGG TGGCACTTTA GTGTTTAGTG GTCGCCCAAC
 2151 ACCACACGCA TACAATCATT TAAATCGCCT AAACGAGCTT GGGCGACCTA
 2201 AGGGCGAAGT GGTTATTGAT GACGATTGGA TCAACCGTAC ATTTAAAGCT
 2251 GAAAACTTCC AAATTAAAGG CGGAAGTACG GTGGTTTCTC GCAATGTTTC
 2301 TTCAATTGAA GGAAATTGGA CAATCAGCAA TAACGCCAAC GCGACATTG
 2351 GTGTTGTGCC AAATCAACAA AATACCATT TGCACGCGTTC AGATTGGACA
 2401 GGATTAACGA CTTGTAAAAC AGTTAATTTA ACCGATAAAA AAGTTATTGA
 2451 TTCCATACCG ACAACACAAA TTAATGGCTC TATTAATTTA ACTAATAATG
 2501 CAACAGTGAA TATTCATGGT TTAGCAAAAC TTAATGGTAA TGTCACTTTA
 2551 ATAAATCATA GCCAATTTAC ATTGAGCAAC AATGCCACCC AAACAGGCAA
 2601 TATCCAACCT TCAAATCACG CAAATGCAAC GGTGGATAAT GCAAACCTGA
 2651 ACGGTAATGT GCATTTAACG GATTCTGCTC AATTTTCTTT AAAAAACAGC
 2701 CATTTTTTCG ACCAAATTCA GGGCGACAAA GACACAACAG TGACGTTGGA
 2751 AAATGCGACT TGGACAATGC CTAGCGATAC TACATTGCAG AATTTAACGC
 2801 TAAATAATAG TACTGTTACG TTAAATTCAG CTTATTCAGC TAGCTCAAAT
 2851 AATGCGCCAC GTCACCGCCG TTCATTAGAG ACGGAAACAA CGCCAACATC
 2901 GGAAGAACAT CGTTTCAACA CATTGACAGT AAATGGTAAA TTGAGCGGGC
 2951 AAGGCACATT CCAATTTACT TCATCTTTAT TTGGCTATAA AAGCGATAAA
 3001 ATAAAATTAT CTAATGACGC TGAAGGCGAT TACACATTAG CTGTTTCGCGA
 3051 CACAGGCAAA GAACCTGTGA CCCTTGAGCA ATTAACTTTA ATTGAAGGCT
 3101 TGGATAATCA ACCCTTGCCA GATAAGCTAA AAATTACTTT AAAAAATAAA
 3151 CACGTTGATG CGGGTGCATG GCGTTATGAA TTAGTGAAGA AAAACGGCGA
 3201 ATTCCGCTTG CATAATCCAA TAAAAGAGCA GGAATTGCGC AATGATTTAG
 3251 TAAAAGCAGA GCAAGTAGAA CGAGCATTAG AAGCAAAACA AGCTGAACTG
 3301 ACTACTAAAA AACAAAAAAC TGAGGCTAAA GTGCGGTCAA AAAGAGCGGC
 3351 GTTTTCTGAT ACCCCGCCTG ATCAAAGCCA GTTAAACGCA TTACAAGCCG
 3401 AACTCGAGAC GATTAATGCC CAACAGCAAG TGGCACAAGC GGTGCAAAAT
 3451 CAGAAAGTAA CTGCACTTAA CCAAAGAAC GAGCAAGTTA AAACCACTCA
 3501 AGATAAAGCA AATTTAGTCT TGGCAACTGC ATTGGTGGAA AAAGAAACCG
 3551 CTCAGATTGA TTTTGCTAAT GCAAAATTAG CTCAGTTGAA TTTAACACAA
 3601 CAACTAGAAA AAGCCTTAGC AGTGGCTGAG CAAGCAGAAA AAGAGCGTAA
 3651 AGCTCAAGAG CAAGCGAAAA GACAACGCAA ACAAAAAGAC TTGATCAGCC
 3701 GTTATTCAAA TAGTGCGTTA TCAGAATTAT CTGCAACAGT AAATAGTATG
 3751 CTTTCCGTTT AAGATGAATT AGATCGTCTT TTTGTAGATC AAGCTCAATC
 3801 TGCGGTGTGG ACAAATATCT CACAGGATAA AAGACGTTAT GATTCTGATG
 3851 CGTTCCGTGC TTATCAGCAG AAAACGAACT TGCGTCAAAT TGGGGTGCAA
 3901 AAAGCCTTAG CTAACGGACG AATTGGGGCA GTTTTCTCGC ATAGCCGTTT
 3951 AGATAATACT TTTGATGAAC AGGTTAAAAA TCACGCAACA TTAACGATGA
 4001 TGTCGGGTTT TGCCCAATAT CAATGGGGTG ATTTACAATT TGGTGTAAAC
 4051 GTGGGAACGG GAATTAGTGC GAGTAAATG GCTGAAGAAC AAAGCCGAAA
 4101 AATTCATCGA AAAGCGATAA ATTATGGCGT GAATGCAAGT TATTCGTTCC
 4151 ATTTAGGGCA ATTGGGTATT CAGCCTTATT TTGGAGTTAA TCGCTATTTT
 4201 ATTGAACGTA AAAATTATCA ATCTGAGGAA GTGAAAGTGC AAACACCGAG

FIG. 24B

45 / 45

4251 CCTTGCATTT AATCGCTATA ATGCTGGAGT ACGGGTCGAT TATACGTTTA
 4301 CCCCACACAGA GAATATCAGC GTTAAGCCTT ATTTCTTCGT CAATTATGTT
 4351 GATGTTTCAA ACGCTAACGT ACAAACCACT GTAAATCGCG CCGTGTGCA
 4401 ACAACCATTT GGACGTTATT GGCAAAAAGA AGTGGGATTA AAAGCGGAAA
 4451 TTTTACATTT CCAACTTTCT GCTTTTATTT CTAAATCTCA AGGTTCGCAA
 4501 CTCGGTAAAC AGCGAAATAT GGGCGTGAAA TTAGGATATC GTTGGTAAAA
 4551 ATCAACATAA TTTTATTCTA ATAATGGAAC TTTATTTAAT TAAAAGTATC
 4601 TAAGTAGCAC CCTATAGGGG ATTAATTAAG AGGATTTAAT AATGAATTTA
 4651 ACTAAAATTT TACCCGCATT TGCTGCTGCA GTCGTATTAT CTGCTTGTGC
 4701 AAAGGATGCA CCTGAAATGA CAAATCATC TGCGCAAATA GCTGAAATGC
 4751 AAACACTTCC AACAATCACT GATAAAACAG TTGTATATTC TTGCAATAAA
 4801 CAAACTGTGA CTGCAGTGTA TCAATTTG

FIG._24C

Amino acid sequence for NTHi strain 1396B Hap protein (first amino acid to last amino acid):

1 MKKTVFRLNF LTACISLGIV SQAWAGHTYF GIDYQYYRDF AENKGKFTVG
 51 AKNIEVYNKN GNLVGTSMTK APMIDFSVVS RNGVAALVGD QYIVSVAHNV
 101 GYTNVDFGAE GQNPQHRFT YKIVKRNNYK NDQTHPYEKD YHNPRLHKFV
 151 TEATPIDMTS DMNGNKYTDR TKYPERVRIG SGWQFWRNDQ NNGDQVAGAY
 201 HYLTAGNTHN QGGAGGGWSS LSGDVRQAGN YGPIPIAGSS GDSGSPMFIY
 251 DAEKQKWLIN GVLRTGNPWA GTENTFQLVR KSFFDEILEK DLRTSFYSPS
 301 GNGAYTITDK GDGSGIVKQQ TGRPSEVRIG LKDDKLPAEG KDDVYQYQGP
 351 NIYLPRLNNG GNLYFGDQKN GTVTLSNIN QGAGGLYFEG NFTVSSENNA
 401 TWQGAGVHVG EDSTVTWKVN GVENDRLSKI GKGTLHVKAK GENKGSISVG
 451 DGKVILEQQA DDQGNKQAFS EIGLVSGRGT VOLNDDKQFN TDKFYFGFRG
 501 GRDLNLGHSL TFKRIQNTDE GATIVNHNAT TESTVTITGS DTINDNTGDL
 551 TNKRDI AFNG WFGDKDDTKN TGRNLVTYNP LNKDNHFLLS GGTNLKGNIT
 601 QDGGTLVFSG RPTPHAYNHL NRLNELGRP K GEVVIDDDWI NRTFKAENFQ
 651 IKGGSTVVS R NVSSIEGNWT ISNNANATFG VVPNQNTIC TRSDWTGLTT
 701 CKTVNLTDKK VIDSIPPTQI NGSINLTNNA TVNIHGLAKL NGNVTLINHS
 751 QFTLSNNATQ TGNIQLSNHA NATVDNANLN GNVHLTDSAQ FSLKNSHFHSH
 801 QIQGDKD TTV TLENATWTMP SDTTLQNLTL NNSTVTLNSA YSASSNNAPR
 851 HRRSLETETT PTSEEHRFNT LTVNGKLSGQ GTFQFTSSLF GYKSDKIKLS
 901 NDAEGDYTLA VRDTGKEPVT LEQLTLIEGL DNQPLPKLK ITLKNKHVDA
 951 GAWRYELVKK NGEFRLHNPI KEQELRNDLV KAEQVERALE AKQAELTTKK
 1001 QKTEAKVRSK RAAFSDTPPD QSQLNALQAE LETINAQQQV AQAVQNQKVT
 1051 ALNQKNEQVK TTQDKANLVL ATALVEKETA QIDFANAKLA QLNLTQQLEK
 1101 ALAVAEQAEK ERKAQEQAQR QRKQKDLISR YSNSALSLS ATVNSMLSVQ
 1151 DELDRFLVDQ AQSAVWTNIS QDKRRYSDA FRAYQQKTNL RQIGVQKALA
 1201 NGRIGAVFSH SRSDNTFDEQ VKNHATLTMM SGFAQYQWGD LQFGVNVGTG
 1251 ISASKMAEEQ SRKIHRKAIN YGVNASYSFH LGQLGIQPYF GVNRYFIERK
 1301 NYQSEEVKVQ TPLAFNRYN AGVRVDYFT PTENISVKPY FFFVNYVDVSN
 1351 ANVQTTVNRA VLQQPFGRYW QKEVGLKAEI LHFQLSAFIS KSQGSQLGKQ
 1401 RNMGVKLG YR W

FIG._25